



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 136995

TO: Ginny Portner
Location: REM-3B02&3C18
Art Unit: 1645
Sunday, November 14, 2004

Case Serial Number: 09/545199

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Examiner Portner,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Toby Port



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
571-272-2507 Remsen E01 D86

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.



136995

STIC-Biotech/ChemLib

From: Portner, Ginny
Sent: Wednesday, November 03, 2004 2:23 PM
To: STIC-Biotech/ChemLib
Subject: 09/545,199

Please search SEQ ID NO 4, and back translate the amino acid to the corresponding nucleotide sequence. The claims recite 70% identity to a nucleotide sequence that encodes SEQ Id NO 4. Thanks

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02
(571) 272-0862

CRFE

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search
NA Sequence: # _____
AA Sequence :# _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 09:43:28 ; Search time 155 Seconds
(without alignments)
668.857 Million cell updates/sec

Title: US-09-545-199F-4

Perfect score: 1429
Sequence: 1 MAGAKERTIKASVKSTQKI.....ARQASITNELNIVAGAAAI 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match %	Length	ID	Description
1	1429	100.0	289	3	AB44522 Virulence
2	1429	100.0	289	5	ABP54474
3	1132.5	79.3	288	3	AB44579 Virulence
4	1132.5	79.3	288	5	ABP5451 Actinobac
5	1077.5	75.4	288	5	ABP54547
6	1004	70.3	288	6	ABM70561 Phototrab
7	1004	70.3	291	7	ADF07164 Bacterial
8	991	69.3	287	7	ACG98289 Escherich
9	978	68.4	297	7	ABO64099 Klebsiell
10	926.5	64.8	291	7	ABO70334 Pseudomon
11	894	62.6	292	6	AD333607 Acinetoba
12	885	61.9	309	8	ADL05162 M. catarr
13	774	54.2	269	4	ABG24481 Novel hum
14	760	53.2	291	6	ABP77380 N. gonorr
15	642.5	45.0	924	4	ABG13826 Novel hum
16	616	43.1	794	4	ABG25667 Novel hum
17	530.5	37.1	304	7	ADH88017 Enterococ
18	524.5	36.7	289	2	AAW56792 L. lactis
19	524.5	36.7	556	4	ABG18379 Novel hum
20	521.5	36.5	288	4	ABG18379 Novel hum
21	518.5	36.3	293	5	ABP27578 Streptoco
22	510.5	35.7	300	7	ADG97657 E. fasciu
23	508.5	35.6	291	5	ABP27579 Streptoco
24	504.5	35.3	289	2	AAW56795 L. lactis
25	501.5	35.1	292	6	ABU01963 S. pneumo

26	501.5	35.1	292	6	ABP81612	Abp81612 Streptoco
27	501.5	35.1	292	8	ADK46870	Adk46870 Streptoco
28	493	34.5	295	5	ABP39397	Abp39397 Staphyloc
29	489	34.2	290	5	ABB47870	Abb47870 Listeria
30	486	34.0	288	6	ABM71838	Abm71838 Staphyloc
31	464	32.5	272	4	AAAG2359	AAg2359 S. epider
32	461.5	32.3	325	4	AAAB79630	AAb79630 Corynebac
33	461.5	32.3	325	4	AAAG91087	AAg91087 C Glutami
34	460.5	32.2	293	2	AAAW56798	AAw56798 S. thermo
35	454	31.8	365	3	AAAG04932	AAg04932 Arabidops
36	454	31.8	373	3	AAAG04931	AAg04931 Arabidops
37	454	31.8	373	8	ADN74311	Adn74311 Thale cre
38	448	31.4	357	5	ABM73921	ABm73921 DNA clone
39	447.5	31.3	327	7	ABBO8865	ABb08865 Corynebac
40	440	30.8	308	7	ADB74399	Adb74399 Mycobacte
41	416.5	29.1	315	4	AAU61796	AAu61796 Propionib
42	416.5	29.1	315	6	ABM58315	ABm58315 Propionib
43	391	27.4	301	3	AAAG04933	AAg04933 Arabidops
44	355.5	24.9	307	5	ABP65479	ABp65479 Bifidobac
45	350.5	24.5	323	3	AAQ08747	AAq08747 Arabidops

ALIGNMENTS

RESULT 1

AA44522
ID AAB44522 standard; protein; 289 AA.

AC AAB44522;

DT 08-FEB-2001 (first entry)

DE Virulence gene protein #2.

XX Virulence gene; antibacterial; vaccine; bacterial infection; septicemia;
XX bronchopneumonia; rhinitis; wound infection.

OS Pasteurella multocida.

PN WC2000061724-A2.

PD 19-OCT-2000.

PF 06-APR-2000; 2000WO-US009218.

PR 09-APR-1999; 99US-0128689P.

PR 10-SEP-1999; 99US-0153453P.

PR (PHAA) PHARMACIA & UPOJOHN INC.

XX Lowery DE, Fuller TE, Kennedy MJ;

XX WPI; 2000-647422/62.

XX N-PSDB; AAC79582.

XX Attenuated Pasteurellaceae bacteria comprising mutations in virulence
XX genes, useful as a live attenuated vaccine against bacterial infections.

XX Claim 39; Page 68-69; 322pp; English.

XX The family Pasteurellaceae encompasses several pathogens that infect a
XX wide variety of animals. The present invention relates to virulence genes
XX from Pasteurellaceae. The present sequence is a protein encoded by one
XX such virulence gene. The virulence genes of the present invention may be
XX mutated in order to produce an inactive gene. The inactive virulence gene
XX may in turn be used to produce a vaccine, which is useful for treating
XX bacterial infections such as septicemias, bronchopneumonias, rhinitis and
XX wound infections

XX Sequence 289 AA;

Query Match 100.0%; Score 1429; DB 3; Length 289;

Best Local Similarity 100.0%; Pred. No. 5.7e-121;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGAKERTKIASVKSTQKITKAMEMVAASKVRKTQERMSSSRPYSETIRNVISHVSKAT 60
DB 1 MAGAKERTKIASVKSTQKITKAMEMVAASKVRKTQERMSSSRPYSETIRNVISHVSKAT 60
QY 61 IGYKHPLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNMKEWKEKDVSVQSLIGS 120
DB 61 IGYKHPLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNMKEWKEKDVSVQSLIGS 120
QY 121 KSINFFQSLGKILTDQSGIGDTPSVEQLIGSVNSMIDAYKKGEVDVYLVYVKFINTMS 180
DB 121 KSINFFQSLGKILTDQSGIGDTPSVEQLIGSVNSMIDAYKKGEVDVYLVYVKFINTMS 180
QY 181 QKPVLEKILPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
DB 181 QKPVLEKILPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
QY 241 EQAARMVAMKAATDNAGNLINELQLVYNKARQASITNELNEIVAGAAAI 289
DB 241 EQAARMVAMKAATDNAGNLINELQLVYNKARQASITNELNEIVAGAAAI 289

RESULT 2
ABP54474
ID ABP54474 standard; protein; 289 AA.
XX AC ABP54474;
XX DT 06-AUG-2003 (revised)
XX DT 24-JAN-2003 (first entry)
DE Pasteurella multocida atpG protein SEQ ID NO:4.
XX AC Antibacterial; vaccine; gram negative bacterial virulence gene;
XX AC identification; virulence; Pasteurellaceae.
XX OS Pasteurella multocida.
XX PN WO200275507-A2.
XX PD 26-SEP-2002.
XX PF 17-JAN-2002; 2002WO-US001971.
XX PR 15-MAR-2001; 2001US-00809665.
XX PA (PHAA) PHARMACIA & UPJOHN CO.
XX PI Lowery DE, Fuller TE, Kennedy MJ;
XX DR N-PSDB; ABQ83459.
XX DR WPI; 2002-740868/80.
XX PI N-PSDB; ABQ83459.
PT New mutant gram-negative bacteria, useful as vaccines and for identifying
FT new anti-bacterial agents that target virulence genes and their products.
XX FS Claim 36; Page 73-74; 350pp; English.
XX CC The present invention describes a gram-negative bacteria comprising a
CC mutation in a gene, where the mutation results in decreased activity of a
CC gene product encoded by the mutated gene. Also described is a method for
CC producing a gram-negative bacteria mutant or an attenuated
CC Pasteurellaceae bacteria. The mutated genes have antibacterial activity
CC and can be used in vaccines. The gram-negative bacteria or the attenuated
CC Pasteurellaceae bacteria can be used as vaccines in the fields of human
CC medicine or veterinary medicine, and for identifying new antibacterial
CC agents that target the virulence genes and their products. ABQ83458 to
CC ABQ83459 and ABP54473 to ABP54551 represents sequences used in the
CC exemplification of the present invention. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX

SQ Sequence 289 AA;
Query Match 100.0%; Score 1429; DB 5; Length 289;
Best Local Similarity 100.0%; Pred. No. 5.7e-121;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGAKERTKIASVKSTQKITKAMEMVAASKVRKTQERMSSSRPYSETIRNVISHVSKAT 60
DB 1 MAGAKERTKIASVKSTQKITKAMEMVAASKVRKTQERMSSSRPYSETIRNVISHVSKAT 60
QY 61 IGYKHPLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNMKEWKEKDVSVQSLIGS 120
DB 61 IGYKHPLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNMKEWKEKDVSVQSLIGS 120
QY 121 KSINFFQSLGKILTDQSGIGDTPSVEQLIGSVNSMIDAYKKGEVDVYLVYVKFINTMS 180
DB 121 KSINFFQSLGKILTDQSGIGDTPSVEQLIGSVNSMIDAYKKGEVDVYLVYVKFINTMS 180
QY 181 QKPVLEKILPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
DB 181 QKPVLEKILPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
QY 241 EQAARMVAMKAATDNAGNLINELQLVYNKARQASITNELNEIVAGAAAI 289
DB 241 EQAARMVAMKAATDNAGNLINELQLVYNKARQASITNELNEIVAGAAAI 289

RESULT 3
AAB44579
ID AAB44579 standard; protein; 288 AA.
XX AC AAB44579;
XX DT 08-FEB-2001 (first entry)
XX DE Virulence gene protein #59.
XX KW Virulence gene; antibacterial; vaccine; bacterial infection; septicemia;
XX KW bronchopneumonia; rhinitis; wound infection.
XX OS Actinobacillus pleuropneumoniae.
XX PN WO200061724-A2.
XX PD 19-OCT-2000.
XX PF 06-APR-2000; 2000WO-US009218.
XX PR 09-APR-1999; 99US-0128689P.
XX PR 10-SEP-1999; 99US-0153453P.
XX PA (PHAA) PHARMACIA & UPJOHN INC.
XX PI Lowery DE, Fuller TE, Kennedy MJ;
XX DR WPI; 2000-647422/62.
XX DR N-PSDB; AAC79654.
XX PT Attenuated Pasteurellaceae bacteria comprising mutations in virulence
XX genes, useful as a live attenuated vaccine against bacterial infections.
XX FS Claim 39; Page 292; 322pp; English.
XX CC The family Pasteurellaceae encompasses several pathogens that infect a
XX wide variety of animals. The present invention relates to virulence genes
XX from Pasteurellaceae. The present sequence is a protein encoded by one
XX such virulence gene. The virulence genes of the present invention may be
XX mutated in order to produce an inactive gene. The inactive virulence gene
XX may in turn be used to produce a vaccine, which is useful for treating
XX bacterial infections such as septicemias, bronchopneumonias, rhinitis and
XX wound infections
XX SQ Sequence 288 AA;

Query Match 79.3%; Score 1132.5; DB 3; Length 288;
 Best Local Similarity 74.4%; Pred. No. 4.4e-94;
 Matches 215; Conservative 45; Mismatches 28; Indels 1; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASMKRKTQERMSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVKSTOKITKAMEMVATSMKRTQERMAASRPYSETIRKVIISHIAGS 60

QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNEKWKEDVSVQSLIGS 120
 DB 61 IGYKHPFLTERDIKKVGLVIVVSTDRGLCGGLNINLFKATLNEFKTWKDKDVSVELGLVGS 120

QY 121 KSINFFQSLGKILTDOSGIGDTPSVQELIGSVNSMIDAYKKEGVVVYLVNKFINTMS 180
 DB 121 KGVSYFQNLGLNVRQVTVGLGDNPEMERIVGAVNEMINAFRNGEVDVAVYVNFENTMS 180

QY 181 QKPVLEKLIPLDELNDLGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QKPVIAQLLEPLKLDDEL-DTKGSDYIYEPNPQVLLDLSLLVRYLETQVYQAVVDNLAS 239

QY 241 EQAARVMVAKAATDNAGNLINELQLVYNKARQASITNELNEIVAGAAAI 289
 DB 240 EQAARVMVAKAATDNAGTLIDELQLVYNKARQASITNELNEIVAGAAAI 288

RESULT 4
 ABP54531
 ID ABP54531 standard; protein; 288 AA.
 AC ABP54531;
 DT 24-JAN-2003 (first entry)
 DE Actinobacillus pleuropneumoniae atpG protein SEQ ID NO:133.
 KW Antibacterial; vaccine; gram negative bacterial virulence gene;
 KW identification; virulence; Pasteurellaceae.
 OS Actinobacillus pleuropneumoniae.
 XX WO200275507-A2.
 PD 26-SEP-2002.
 PF 17-JAN-2002; 2002WO-US001971.
 PR 15-MAR-2001; 2001US-00809665.
 PA (PHAA) PHARMACIA & UPJOHN CO.
 PI Lowery DE, Fuller TE, Kennedy MJ;
 DR WPI; 2002-740868/80.
 XX N-PSDB; ABQ83531.
 PT New mutant gram-negative bacteria, useful as vaccines and for identifying
 PT new anti-bacterial agents that target virulence genes and their products.
 PS Claim 36; Page 298-299; 350pp; English.

The present invention describes a gram-negative bacteria comprising a
 mutation in a gene, where the mutation results in decreased activity of a
 gene product encoded by the mutated gene. Also described is a method for
 producing a gram-negative bacteria mutant or an attenuated
 Pasteurellaceae bacteria. The mutated genes have antibacterial activity
 and can be used in vaccines. The gram-negative bacteria or the attenuated
 Pasteurellaceae bacteria can be used as vaccines in the fields of human
 medicine or veterinary medicine, and for identifying new antibacterial
 agents that target the virulence genes and their products. ABQ83458 to
 ABQ83578 and ABP54473 to ABP54551 represents sequences used in the
 exemplification of the present invention

Sequence 288 AA;
 Query Match 79.3%; Score 1132.5; DB 5; Length 288;
 Best Local Similarity 74.4%; Pred. No. 4.4e-94;
 Matches 215; Conservative 45; Mismatches 28; Indels 1; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASMKRKTQERMSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVKSTOKITKAMEMVATSMKRTQERMAASRPYSETIRKVIISHIAGS 60

QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNEKWKEDVSVQSLIGS 120
 DB 61 IGYKHPFLTERDIKKVGLVIVVSTDRGLCGGLNINLFKATLNEFKTWKDKDVSVELGLVGS 120

QY 121 KSINFFQSLGKILTDOSGIGDTPSVQELIGSVNSMIDAYKKEGVVVYLVNKFINTMS 180
 DB 121 KGVSYFQNLGLNVRQVTVGLGDNPEMERIVGAVNEMINAFRNGEVDVAVYVNFENTMS 180

QY 181 QKPVLEKLIPLDELNDLGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QKPVIAQLLEPLKLDDEL-DTKGSDYIYEPNPQVLLDLSLLVRYLETQVYQAVVDNLAS 239

QY 241 EQAARVMVAKAATDNAGNLINELQLVYNKARQASITNELNEIVAGAAAI 289
 DB 240 EQAARVMVAKAATDNAGTLIDELQLVYNKARQASITNELNEIVAGAAAI 288

RESULT 5
 ABP54547
 ID ABP54547 standard; protein; 288 AA.
 AC ABP54547;
 DT 29-AUG-2003 (revised)
 DT 24-JAN-2003 (first entry)
 DE Pasteurella haemolytica atpG protein SEQ ID NO:167.
 KW Antibacterial; vaccine; gram negative bacterial virulence gene;
 KW identification; virulence; Pasteurellaceae.
 OS Mannheimia haemolytica.
 XX WO200275507-A2.
 PD 26-SEP-2002.
 PF 17-JAN-2002; 2002WO-US001971.
 PR 15-MAR-2001; 2001US-00809665.
 PA (PHAA) PHARMACIA & UPJOHN CO.
 PI Lowery DE, Fuller TE, Kennedy MJ;
 DR WPI; 2002-740868/80.
 XX N-PSDB; ABQ83549.
 PT New mutant gram-negative bacteria, useful as vaccines and for identifying
 PT new anti-bacterial agents that target virulence genes and their products.
 PS Claim 36; Page 330-331; 350pp; English.

The present invention describes a gram-negative bacteria comprising a
 mutation in a gene, where the mutation results in decreased activity of a
 gene product encoded by the mutated gene. Also described is a method for
 producing a gram-negative bacteria mutant or an attenuated
 Pasteurellaceae bacteria. The mutated genes have antibacterial activity
 and can be used in vaccines. The gram-negative bacteria or the attenuated
 Pasteurellaceae bacteria can be used as vaccines in the fields of human
 medicine or veterinary medicine, and for identifying new antibacterial
 agents that target the virulence genes and their products. ABQ83458 to
 ABQ83578 and ABP54473 to ABP54551 represents sequences used in the

CC exemplification of the present invention. (Updated on 29-AUG-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 288 AA;
 Query Match 75.4%; Score 1077.5; DB 5; Length 288;
 Best Local Similarity 72.0%; Pred. No. 4.3e-89;
 Matches 208; Conservative 46; Mismatches 34; Indels 1; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTOKITKAMEWVAASXKRTQERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVRNTOKITKAMEWVAASXKRTQERMAASRPAESIRKASHIAKGN 60
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLFTVLNEMKEWKEKDVSVQLSLIGS 120
 DB 61 LEYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLFTVLNEMKEWKEKDVSVQLSLIGS 120
 QY 121 KSINFFQSLGKILTPQDSIGDTPSVQELIGSVNSMIDAYKGEVDVYVLYVYNKFINIMS 180
 DB 121 KGISFFNPMGLEIKGHINGLGTDPAMEDLVGVNMGVMVAYREGEIDEVYVYVYNRFINIMS 180
 QY 181 OKPVLKLIPLPELDNDELGERKQVWDYIYEPDAKVLLDNLVLYLESQVYQAAVENLAS 240
 DB 181 OKPTVQQLPLPALNDSGL-EQTSWDYLYEPDPKALLDNLVLYLESQVYQAAVDNLAS 239
 QY 241 EQAARWVAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
 DB 240 EQAARWVAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 288
 RESULT 5
 ARM70561
 ID ARM70561 standard; protein; 288 AA.
 AC ARM70561;
 DT 20-NOV-2003 (first entry)
 DE Photorhabdus luminescens protein sequence #3658.
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 OS Photorhabdus luminescens.
 PN WO200294867-A2.
 XX
 XX 28-NOV-2002.
 XX 07-FEB-2002; 2002WO-IB003040.
 XX 07-FEB-2001; 2001FR-00001659.
 XX (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 XX Buchrieser C;
 XX WPI; 2003-148459/14.
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX Claim 2; SEQ ID NO 3658; 1205pp; French.
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the

CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins
 XX
 SQ Sequence 288 AA;
 Query Match 70.3%; Score 1004; DB 6; Length 288;
 Best Local Similarity 67.5%; Pred. No. 2e-82;
 Matches 195; Conservative 48; Mismatches 44; Indels 2; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTOKITKAMEWVAASXKRTQERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVQNTOKITKAMEWVAASXKRTQERMAASRPAETIRSVIGHALGN 60
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLFTVLNEMKEWKEKDVSVQLSLIGS 120
 DB 61 LEYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLFTVLNEMKEWKEKDVSVQLSLIGS 120
 QY 121 KSINFFQSLGKILTPQDSIGDTPSVQELIGSVNSMIDAYKGEVDVYVLYVYNKFINIMS 180
 DB 121 KATSFASVGGNVVAVQVTGNGNPSLSLIGVNIIMLRAYDEGLDKLYVYVYNKFINIMS 180
 QY 181 OKPVLKLIPLPELDNDELGERKQVWDYIYEPDAKVLLDNLVLYLESQVYQAAVENLAS 240
 DB 181 OKPTVQQLPLPALNDSGL-EQTSWDYLYEPDPKALLDNLVLYLESQVYQAAVENLAS 239
 QY 241 EQAARWVAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
 DB 239 EQAARWVAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 287
 RESULT 7
 ADF07164
 ID ADF07164 standard; protein; 291 AA.
 AC ADF07164;
 DT 12-FEB-2004 (first entry)
 DE Bacterial polypeptide #3277.
 KW Proteus mirabilis infection; bacterial infection; antibacterial;
 KW immunostimulant.
 OS Proteus mirabilis.
 PN US6605709-B1.
 XX 12-AUG-2003.
 XX 05-APR-2000; 2000US-00543681.
 XX 09-APR-1999; 99US-0128706P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Breton GL;
 XX WPI; 2003-895291/82.
 XX N-PSDB; ADF02992.

XX New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX Disclosure; SEQ ID NO 7449; 870pp; English.
PS
XX The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX
SQ Sequence 291 AA;

Query Match 70.3%; Score 1004; DB 7; Length 291;
Best Local Similarity 67.1%; Pred. No. 2e-82;
Matches 194; Conservative 49; Mismatches 44; Indels 2; Gaps 1;
QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASRKQTKQERMSSRPYSSETIRNVISHVSKAT 60
DB 5 MAGAKEIRSKIASVQNTQKITKAMEMVAASRKQTKQERMSSRPYSSETIRNVISHVSKAT 64
QY 61 IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTVLNMKWKEDVSVQSLIGS 120
DB 65 LEYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTVLNMKWKEDVSVQSLIGS 124
QY 121 KSNFFQSLGKILQDSGIGDTPSVQELIGSVNMDAYKKGVDVYVLYVYKFNITMS 180
DB 125 KAVSFFASVGGNVVQVVTGMDDPQLSDLGPNVIMLQAYDEGLDKLYVYKFNITMA 184
QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVYLESQVYQAAVENLAS 240
DB 185 QEPKILQVLPFGDDDELKES--WDYLYEPDPKTLTLLRYYTESQVYQSVVENLAS 242
QY 241 EQAARWAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
DB 243 EQAARWAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAV 291

RESULT 8
AAG98289
ID AAG98289 standard; protein; 287 AA.
AC AAG98289;
XX
XX 21-SEP-2001 (first entry)
XX Escherichia coli protein sequence SEQ ID NO:337.
XX Escherichia coli; identification; proliferation; microorganism;
XX antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;
XX bacterial growth inhibition.
XX Escherichia coli.
XX WO200148209-A2.
XX
XX 05-JUL-2001.
XX 19-DEC-2000; 2000WO-US034419.
XX 23-DEC-1999; 99US-0173005P.
XX (ELIT-) ELITRA PHARM INC.

XX
PI Forsyth RA, Ohlsen KL, Zyskind JW;
XX WPI: 2001-457376/49.
DR N-PSDB; AAH81345.
XX
PT Novel nucleic acids encoding proteins required for Escherichia coli
PT proliferation, useful for screening for antimicrobial agents.
XX
XX Claim 19; Page 468; 596pp; English.

XX The present invention describes a purified or isolated nucleic acid
CC sequence (I) consisting essentially of one of the 93 nucleotide sequences
CC given in AAH81202 to AAH81294, where expression of the nucleic acid in a
CC microorganism is capable of inhibiting proliferation of a microorganism.
CC (II) have antibacterial and antibiotic activities, and can be used in gene
CC therapy. Expression of (I) in a microorganism inhibits proliferation of
CC the microorganism, and the manufactured antibiotic is useful for reducing
CC the activity or level of a gene product required for proliferation of a
CC microorganism in a subject, specifically humans. The nucleic acids that
CC inhibit bacterial growth or proliferation can be used as antisense
CC therapeutics for killing bacteria. In addition to therapeutic
CC applications, the nucleic acid sequences complementary to sequences
CC required for proliferation can be used as diagnostic tools. For example,
CC nucleic acid probes complementary to proliferation-required sequences
CC that are specific for particular species of microorganisms can be used as
CC probes to identify particular microorganism species in clinical
CC specimens. AAH81295 to AAH81487 encode the Escherichia coli proteins
CC given in AAG98239 to AAG98431, and AAH81488 to AAH81491 represent
CC oligonucleotides, which are used in the exemplification of the present
CC invention
XX
SQ Sequence 287 AA;

Query Match 69.3%; Score 991; DB 4; Length 287;
Best Local Similarity 56.1%; Pred. No. 3e-81;
Matches 191; Conservative 47; Mismatches 49; Indels 2; Gaps 1;
QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASRKQTKQERMSSRPYSSETIRNVISHVSKAT 60
DB 1 MAGAKEIRSKIASVQNTQKITKAMEMVAASRKQTKQERMSSRPYSSETIRNVISHVSKAT 60
QY 61 IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTVLNMKWKEDVSVQSLIGS 120
DB 61 LEYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTVLNMKWKEDVSVQSLIGS 120
QY 121 KSNFFQSLGKILQDSGIGDTPSVQELIGSVNMDAYKKGVDVYVLYVYKFNITMS 180
DB 121 KGVSFNSVGGNVVQVVTGMDDPQLSDLGPNVIMLQAYDEGLDKLYVYKFNITMS 180
QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVYLESQVYQAAVENLAS 240
DB 181 QVPTISQLPLPASDDDL--KHKSNDYLYEPDPKALDTLLRYYTESQVYQSVVENLAS 238
QY 241 EQAARWAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
DB 239 EQAARWAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAV 287

RESULT 9
ABO64099
ID ABO64099 standard; protein; 297 AA.
XX
XX ABO64099;
AC
XX
XX 29-JUL-2004 (first entry)
XX Klebsiella pneumoniae polypeptide seqid 10616.
XX
XX Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
XX Klebsiella pneumoniae.

```

XX PN US6610836-B1.
XX PD 26-AUG-2003.
XX PF 27-JAN-2000; 2000US-00489039.
XX PR 29-JAN-1999; 99US-0117747P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton GL, Osborne M;
XX PD N-PSDB; ACH97650.
XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX PT preparing a vaccine composition against Klebsiella pneumoniae.
XX PS Disclosure; SEQ ID NO 10616; 932pp; English.
XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
XX CC pneumoniae polypeptide. Also described are: a recombinant expression
XX CC vector comprising the nucleic acid, operably linked to a transcription
XX CC regulatory element; and a cell comprising the recombinant expression
XX CC vector. The nucleic acid is useful for preparing a vaccine composition
XX CC against Klebsiella pneumoniae. This is the amino acid sequence of a
XX CC Klebsiella pneumoniae polypeptide of the invention
XX SQ Sequence 297 AA;
    Query Match      68.4%; Score 978; DB 7; Length 297;
    Best Local Similarity 65.1%; Pred. No. 4.7e-80;
    Matches 188; Conservative 50; Mismatches 49; Indels 2; Gaps 1;
QY 1 MAGAKEIRTKIASVSKTQKITKAMEMVAASQKRTQERMSSSRPYSETIRNVISHVSKAT 60
DB 11 MAGAKEIRSKIASVNTQKITKAMEMVAASQKRSQERMAASRPYADTKRKVIGHLANGN 70
QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 120
DB 71 LEYKHPFLVDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 130
QY 121 KSNFFQSLGKILQDQSGIGDTPSVBOLIGSVNSMIDAYKKGEVDVYLVYVYVYVYVYV 180
DB 131 KGVFFNSVGGNVVAQVIGMDNFSLELGPVKVWLQAYDEGRDLKLYVVSNNKFINTMS 190
QY 191 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
DB 191 QVPTITQLLPASEDADL--KRSWDYLYEPDPKALDITLLRRYVESQVYQGVVENLAS 248
QY 241 EQAARMVAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAI 289
DB 249 EQAARMVAMKAATDNGSLIKELQVYNKARQASITQELTEIVSGAAAV 297

RESULT 10
ABO70334
ID ABO70334 standard; protein; 291 AA.
XX AC ABO70334;
XX AC ABO70334;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #2509.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.

```

```

XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX PD N-PSDB; ABD03905.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 19080; 455pp; English.
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using biochip technology. Sequences ABO67826-
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 291 AA;
    Query Match      64.8%; Score 926.5; DB 7; Length 291;
    Best Local Similarity 62.3%; Pred. No. 2.1e-75;
    Matches 180; Conservative 50; Mismatches 56; Indels 3; Gaps 1;
QY 1 MAGAKEIRTKIASVSKTQKITKAMEMVAASQKRTQERMSSSRPYSETIRNVISHVSKAT 60
DB 6 MAGAKEIRSKIASIKSTQKITNAMEKVAVKVKAQMEVYACRPYAEIRIQVIGHLANAN 65
QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 120
DB 66 PEYRHFPMVEREVRVGYIYVSSDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 125
QY 121 KSNFFQSLGKILQDQSGIGDTPSVBOLIGSVNSMIDAYKKGEVDVYLVYVYVYVYVYV 180
DB 126 KGASFFRSFGNVAAISHLGEPSINDLIGSVKVMYDAYLEGIDRLFFVSNKFVNTMT 185
QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
DB 186 QKPTVEQLIPLVADDQDEL---KHHWDYLYEPDAKSLDGLLVRYVESQVYQAAVENNAC 242
QY 241 EQAARMVAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAI 289
DB 243 EQAARMVAMKAATDNAGELISDLOIYNKARQAITQELTEIVSGAAAV 291

RESULT 11
ADA33607
ID ADA33607 standard; protein; 292 AA.
XX AC ADA33607;
XX AC ADA33607;
XX DT 20-NOV-2003 (first entry)
XX DE Acinetobacter baumannii protein #768.

```

```

XX Moraxella catarrhalis; infection.
XX Moraxella catarrhalis.
XX US6673910-B1.
XX 06-JAN-2004.
XX 04-APR-2000; 2000US-00540236.
XX 08-APR-1999; 99US-0128416P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Breton GL;
XX WPI; 2004-178127/17.
XX N-PSDB; ADL03242.
XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
XX preparing a composition for diagnosing, preventing or treating infection
XX caused by Moraxella catarrhalis.
XX Disclosure; SEQ ID NO 2848; 429pp; English.
XX The invention relates to an isolated nucleic acid encoding an Moraxella
XX catarrhalis polypeptide. The nucleic acid is useful for preparing a
XX composition for diagnosing, preventing or treating infection caused by
XX Moraxella catarrhalis. The present sequence represents the amino acid
XX sequence of a M. catarrhalis protein.
XX Sequence 309 AA;

Query Match 61.9%; Score 885; DB 8; Length 309;
Best Local Similarity 56.7%; Pred. No. 1.3e-71;
Matches 164; Conservative 63; Mismatches 62; Indels 0; Gaps 0;

QY 1 MAGAKEIRTKIASVKSTOKITKAMVWVAASRKQTKQRMSSSRPSETIRNVISHVSKAT 60
Db 20 MASKEIRAKTSIKSTOKITRAMQWVAASRKQRAQRMELGRPSYSDGIRRVISHLVQAA 79
QY 61 IGYKHPFLVDREVKKVGMIVYSTDRGLCGGLNVNLFKTVLNEMKEWKEDVSVQLSLIGS 120
Db 80 SDYKHPYMINRPVNRGVVVVTSRDLGAGGLNINLFFKLLKTVKSVQESVIEFAVIGA 139
QY 121 KSINFFQSLGKILTDQSGIGDTPSVQELIGSVNSMIDAYKKGEDVVVLYVYNKFTNWS 180
Db 140 KGVFFKNGFGRVTSVAVTDYGNPALEQINTPVQTMDDYLANGKLDRIYLVYNQFINAMA 199
QY 181 QKPVLEKLIPLPDLNDELGERQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
Db 200 QKPVEQIVPLASEEDDTLOAHSDYIYEPDTKTLIDSLRLVIESVYQSVRENIAS 259
QY 241 EQARVWVMAKAATDNAGNLINELQVYNKAROASITNELNEIVAGAAAI 289
Db 260 EQSARVWVMAKAATDNAGNLIKDLQVYNKAROAITREISEIVGGAANV 308

RESULT 13
ABG24481
ID ABG24481 standard; protein; 969 AA.
XX AC ABG24481;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #24472.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.

XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
XX plant biocontrol agent.
XX Acinetobacter baumannii.
XX US6562958-B1.
XX 13-MAY-2003.
XX 04-JUN-1999; 99US-00328352.
XX 08-JUN-1998; 98US-0088701P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Breton G, Bush D;
XX WPI; 2003-576092/54.
XX N-PSDB; ADA29481.
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX for diagnosing a bacterial disease, as components of antibacterial
XX vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX plants.
XX Example; SEQ ID NO 4894; 328pp; English.
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX The A. baumannii nucleic acids and polypeptides are useful as reagents
XX for diagnosing a bacterial disease, as components of antibacterial
XX vaccines, as targets for antibacterial drugs, to detect the presence of
XX A. baumannii and other Acinetobacter species in a sample, in screening
XX compounds for the ability to interfere with the A. baumannii life cycle
XX or to inhibit A. baumannii infection, and as biocontrol agents for
XX plants. The present sequence represents the amino acid sequence of an A.
XX baumannii protein.
XX Sequence 292 AA;

Query Match 62.6%; Score 894; DB 6; Length 292;
Best Local Similarity 59.7%; Pred. No. 1.9e-72;
Matches 173; Conservative 56; Mismatches 59; Indels 2; Gaps 2;

QY 1 MAGAKEIRTKIASVKSTOKITKAMVWVAASRKQTKQRMSSSRPSETIRNVISHVSKAT 60
Db 4 NANKEIRAKVASIKSTOKITRAMQWVAASRKQRAQRMELGRPSYSDGIRRVISHLVQAA 63
QY 61 IGYKHPFLVDREVKKVGMIVYSTDRGLCGGLNVNLFKTVLNEMKEWKEDVSVQLSLIGS 120
Db 64 PEYKHYVWDRPVKRVGVIIIVSSDRGLAGGLNINLFFKVVQHVKAQEQSIEVQFALIGQ 123
QY 121 KSINFFQSLGKILTDQSGIGDTPSVQELIGSVNSMIDAYKKGEDVVVLYVYNKFTNWS 180
Db 124 KAVSFKNYGGKVLGAATTQIGDAPSLQLTGSQVQWMLDAFDKGLDRIYLVNSGPNVAMT 183
QY 181 QKPVLEKLIPLPDLNDELGERQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLA 239
Db 184 QKEVQEVLPVLAPEEGDDL-NRYTGWYDIYEPAEELNGLLVRIEISVYQGVVENVA 242
QY 240 EQARVWVMAKAATDNAGNLINELQVYNKAROASITNELNEIVAGAAAI 289
Db 243 EQSARVWVMAKAATDNAGNLIKDLQVYNKAROAITQEISEIVGGAANV 292

RESULT 12
ADL05162
ID ADL05162 standard; protein; 309 AA.
XX AC ADL05162;
XX DT 06-MAY-2004 (first entry)
XX DE M. catarrhalis protein #928.

```

```
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI WPI; 2001-639362/73.
XX DR N-PSDB; AAS88666.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 54840; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological actions. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 969 AA;
Query Match 54.2%; Score 774; DB 4; Length 969;
Best Local Similarity 55.2%; Pred. No. 7.9e-61;
Matches 155; Conservative 48; Mismatches 70; Indels 8; Gaps 3;
QY 7 IRTKIASVSTQKITKAMEMVAASKVRK--TQRMSSSRPYSETIRNVISHVSKATGYK 64
Db 672 IRTSCTERTGTRWPGRLLESDSFPHKLSQE----SYPAETMRKVIGHLAHNGLEYK 727
QY 65 HPFLVDREVKKYGMIVSTDRGLCGGLNVNLFKTVLNEMKEWKEKDVSVQLSLIGSKIN 124
Db 728 HPVLEDRDVKRVGYLVSTDRGLCGGLNINLFKLLAEMKWTBKGVOCDLWIGSKGSV 787
QY 125 FOSLGIGIKILTQDSGIGDTPSPVEQLIGSVNSMIDAYKKGEVDVYLVNKFINTMSQKV 184
Db 788 FNSVGGNVVAQVTGMDGNPSLSLIGSVKVMQAYDEVRLDKLYICNHQFITMSQVPT 847
QY 185 LEKLIPLELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLASEQAA 244
Db 848 ISQLPLPASDDDDJ--KHKSWDYIYEPDPKALLDTLLRRVYESQVYQGVVENLASEQAA 905
QY 245 RVAMVKAATDNAGNINELQLVYNKARQASITNELNEIVAG 285
Db 906 RVAMVKAATDNGSLIKELQLVYNKARQASITQELTEIVSG 946
```

```
RESULT 14
ABP77380
XX ID ABP77380 standard; protein; 291 AA.
XX AC ABP77380;
XX DT 07-MAR-2003 (first entry)
XX DE N. gonorrhoeae amino acid sequence SEQ ID 1290.
XX KW Antibacterial; infection; vaccine; gene therapy.
XX OS Neisseria gonorrhoeae.
XX FN WO200279243-A2.
XX PD 10-OCT-2002.
XX PF 12-FEB-2002; 2002WO-IB002069.
XX PR 12-FEB-2001; 2001GB-00003424.
XX PA (CHIR-) CHIRON SPÅ.
XX PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX PI WPI; 2003-058415/05.
XX DR N-PSDB; AB238350.
XX CC New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX CC medicament for treating or preventing N. gonorrhoeae infection.
XX PS Disclosure; Page 281; 815pp; English.
XX CC The present invention relates to proteins from Neisseria gonorrhoeae.
XX CC Also disclosed are the nucleic acid molecules encoding the proteins and
XX CC antibodies that specifically bind to the proteins. The composition
XX CC comprising the protein, nucleic acid or antibody is useful for the
XX CC manufacture of a medicament for treating or preventing N. gonorrhoeae
XX CC infection, this may be in the form of a vaccine or gene therapy.
XX CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
XX CC molecules of the invention
XX SQ Sequence 291 AA;
Query Match 53.2%; Score 760; DB 6; Length 291;
Best Local Similarity 51.5%; Pred. No. 2.7e-60;
Matches 150; Conservative 64; Mismatches 75; Indels 2; Gaps 2;
QY 1 MAGAKEIKTKIASVKSTQKITKAMEMVAASKVRKKTQERMSSSRPYSETIRNVISHVSKAT 60
Db 1 MAVGKEILTKIRSVQNTQKITKAMQVSTSKVRKTQERMSSSRPYSETIRNVISHVSKAT 60
QY 61 IGYKHPELVLD-REVKKYGMIVVSTDRGLCGGLNVNLFKTVLNEMKEWKEKDVSVQLSLIG 119
Db 61 TDHGTPLESREIRRVGFILITSDKGLCGGLNANLVKFLAQVQYRNOGIEEBEIVCLG 120
QY 120 SKSINFFQSLGKILTDSDGIGDTPSPVEQLIGSVNSMIDAYKKGEVDVYLVNKFINTM 179
Db 121 SKGLMACQSIGLVNVASAVNLGDTPKVEMLLGPTLTFQRYEKHEIDRIHLVYSGFVNTM 180
QY 180 SQKPVLEKILPELP-LDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENL 238
Db 181 RQERMEVLLPIGENVIGDSAPKSPFSWEYRYEPTALVLELVRRYLESVYQALSNDM 240
QY 239 ASEQAARVAMKAATDNAGNINELQLVYNKARQASITNELNEIVAGAAAI 289
Db 241 ASEQAARVAMKAATDNAGNINELQLVYNKARQASITNELSEIVAGAAV 291
RESULT 15
ABG13826
XX ID ABG13826 standard; protein; 924 AA.
```

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 09:46:43 ; Search time 16 seconds
(without alignments)
1737.915 Million cell updates/sec

Title: US-09-545-199f-4

Perfect score: 1429

Sequence: 1 MAGAKEIRTKIASVKSTQKI.....ARQASITNELNEIVAGAAAI 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

PIR79:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1173	82.1	289	1 E64071	H+-transporting tw
2	998.5	69.9	288	2 G82036	ATP synthase Fl, g
3	991	69.3	287	1 PWECG	H+-transporting tw
4	991	69.3	287	2 C91213	membrane-bound ATP
5	991	69.3	287	2 D86059	membrane-bound ATP
6	989	69.2	287	2 AF0500	H+-transporting tw
7	978	68.4	287	2 AB0954	ATP synthase gamma
8	976.5	68.3	288	1 S06081	H+-transporting tw
9	926.5	64.8	286	2 D82952	ATP synthase gamma
10	808	56.5	290	2 G84930	H+-transporting tw
11	795	55.6	287	2 H82715	ATP synthase gamma
12	754	52.8	291	2 B81970	H+-transporting tw
13	753	52.7	291	2 H81024	ATP synthase gamma
14	566.5	39.6	285	2 G31482	H+-transporting tw
15	536.5	37.5	285	2 C84119	ATP synthase gamma
16	533.5	37.3	287	2 C69592	H+-transporting tw
17	523.5	36.6	289	2 E86845	H+-transporting tw
18	519.5	36.4	289	2 S17725	H+-transporting tw
19	509.5	35.7	315	2 AD1807	ATP synthase chain
20	508.5	35.6	315	2 H31050	H+-transporting tw
21	505.5	35.4	286	2 S01402	H+-transporting tw
22	504.5	35.3	314	1 FWBYG	H+-transporting tw
23	503.5	35.2	316	1 PWYCG	H+-transporting tw
24	501.5	35.1	292	2 G95175	ATP synthase Fl, g
25	501.5	35.1	292	2 H98041	H+-transporting tw
26	496.5	34.7	282	2 S28040	Na+-transporting A
27	495.5	34.7	315	2 S36979	H+-transporting tw
28	494.5	34.6	315	2 S32401	H+-transporting tw
29	492	34.4	290	2 AB1391	H+-transporting At

30	492	34.4	290	2 AD1766	H+-transporting At
31	490	34.3	288	2 E90003	ATP synthase gamma
32	487	34.1	292	2 JC5740	membrane-bound pro
33	477.5	33.4	311	2 S49845	H+-transporting tw
34	476.5	33.3	305	2 A70775	probable atpG prot
35	471.5	33.0	301	2 E64661	H+-transporting tw
36	468	32.8	299	1 FQQFG	H+-transporting tw
37	466.5	32.6	301	2 E71855	ATP synthase Fl, c
38	460	32.2	294	2 B81427	H+-transporting tw
39	459.5	32.2	292	2 G97673	ATP synthase gamma
40	459.5	32.2	292	2 AF2898	ATP synthase gamma
41	458	32.1	377	1 PNTWG	H+-transporting tw
42	455	31.8	364	1 FWSFG	H+-transporting tw
43	454	31.8	373	2 B39732	H+-transporting tw
44	449.5	31.5	286	2 S04673	H+-transporting tw
45	449.5	31.5	386	2 A39732	H+-transporting tw

ALIGNMENTS

RESULT 1

E64071

H+-transporting two-sector ATPase (EC 3.6.3.14) gamma chain - Haemophilus influenzae (st C;Species: Haemophilus influenzae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: E64071
R;Pfleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Maidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: E64071
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-289 <TIGR>

Query Match 82.1%; Score 1173; DB 1; Length 289;

Best Local Similarity 76.5%; Pred. No. 1.5e-69;

Matches 221; Conservative 42; Mismatches 26; Indels 0; Gaps 0;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEVVAASKMKTKTOERSSSRPYSETIRNVISHVSKAT 60

Db 1 MAGAKEIRTKIASVQSTQKITKAMEVVAASKMKTKTOERSSSRPYSETIRNVISHVSKAS 60

QY 61 IGYKHPFLVDREVKKVGMIVSTDRGCGGLNVNLFKTVLNMKEWKEKDVSVQLSLIGS 120

Db 61 IGYKHPFLVEREYKKGILVISTDRGCGGLNVNLFKTVLNMKEWKEKDVSVQLSLIGS 120

QY 121 KSNFFQSLGKILITDSDGIGDTPSVEQLICSVNMDIAYKKGEDVDVYLVYNKFINMTS 180

Db 121 KGSFRRSGFNKKGSLGSDTPALEELIGVANTFDAYRNGEIDAVYIAYNKFVNTMS 180

QY 181 QKPVLEKLPLPELDNDELGERKQVWDYIYEPDAKVLNLDNLVRYLESQVYQAAVENLAS 240

Db 181 QKPVQQLVPLPESKDDHLNEQQTWDYLYEPKPKVLDLSLVYLESQVYQAAVENLAS 240

QY 241 EQAARVMVAKATDAGNLINELQVYVYKARQASITNELNEIVAGAAAI 289

Db 241 EQAARVMVAKATDAGNLINELQVYVYKARQASITNELNEIVAGAAAI 289

RESULT 2

G82036

ATP synthase Fl, gamma chain VC2765 [imported] - Vibrio cholerae (strain N16961 serogrou C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82036

R; Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B. 1., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A/Reference number: A82035; MUID:20406833; PMID:10952301
 A/Accession: G82036
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-288 <HEI>
 A/Cross-references: UNIPROT:Q9K0H4; GB:AE004342; GB:AE003852; NID:G9657359; PIDN:AAF9590
 A/Experimental source: serogroup O1; strain N16961; biotype El Tor
 C/Genetics:
 A/Gene: VC2765
 A/Map position: 1
 C/Superfamily: H+-transporting ATP synthase gamma chain

Query Match 69.9%; Score 998.5; DB 2; Length 288;
 Best Local Similarity 65.1%; Pred. No. 3.7e-58;
 Matches 188; Conservative 54; Mismatches 46; Indels 1; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASQWRKTOERMSSSRPYSPTIRNVISHVSKAT 60
 Db 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASQWRKTOERMSSSRPYSPTIRNVISHVSKAT 60

QY 61 IYKHPFLVDREVKKVGMIVSTDRGLCGGLNVLNFKTLNEMKEWKEKDVSVQSLIGS 120
 Db 61 LEYRHPYLEERAKRQVGIITDRGLCGGLNVLNFKTLNEMKEWKEKDVSVQSLIGS 120

QY 121 KSINFPQSLGKILITQDSIGDTPSVQGLGSVNSMIDAYKGEVDVYLVYKFNFTMS 180
 Db 121 KATAPFNNSGAKVAQVSGLDGSLDELIGSVGMUKYKDGELRLYLVFNQFNTMV 180

QY 181 QKPVLEKLIPLPELDNDELGERKQWVDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 Db 181 QKPIDQLPLPKDSDDM-QRDHNDYIYEPDPKPLDALLRFTESQVYQAAVENLAS 240

QY 241 EQARWVAMKATDNAGNLNELQVYNKARQASITNELNEIVAGAAAI 289
 Db 240 EQARWVAMKATDNASNLDDQLQVYNKARQASITNELNEIVAGAAAV 288

RESULT 3
 PNEG
 H+-transporting two-sector ATPase (EC 3.6.3.14) gamma chain - *Escherichia coli* (strain K
 C/Species: *Escherichia coli*
 C/Date: 02-Apr-1982; #sequence, revision 15-Oct-1982; #text, change 09-Jul-2004
 C/Accession: A01038; I41276; A01039; I55328; A53160; F65176
 C/Sarasté, M.; Gay, N.J.; Eberle, A.; Runswick, M.J.; Walker, J.E.
 Nucleic Acids Res. 9, 5287-5296, 1981
 A>Title: The *atp* operon: nucleotide sequence of the genes for the gamma, beta, and epsilon
 A/Reference number: A93742; MUID:82059507; PMID:6272217
 A/Accession: A01038
 A/Molecule type: DNA
 A/Residues: 1-287 <SAR>
 A/Cross-references: UNIPROT:P00837; GB:V00267; NID:G41036; PIDN:CAA23526.1; PID:G41037
 R; Kanazawa, H.; Futai, M.
 Ann. N. Y. Acad. Sci. 402, 45-64, 1982
 A>Title: Structure and function of H+-ATPase: What we have learned from *Escherichia coli*
 A/Reference number: I41271; MUID:83176724; PMID:6303339
 A/Accession: I41276
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-287 <RES>
 A/Cross-references: GB:M25464; NID:G146318; PIDN:AAA83874.1; PID:G146324
 R; Kanazawa, H.; Kayano, T.; Mabuchi, K.; Futai, M.
 Biochem. Biophys. Res. Commun. 103, 604-612, 1981
 A>Title: Nucleotide sequence of the genes coding for alpha, beta and gamma subunits of b
 A/Reference number: A90103; MUID:82134798; PMID:6277310
 A/Accession: A01039
 A/Molecule type: DNA
 A/Residues: 1-37; 'AWRPVLMQKPGAQ', 52-82, 'I', 84-92, 'T', 94-111, 'HSTSGAAR', 121, 'D', 123-127
 A/Cross-references: GB:V00312; NID:G42282; PIDN:CAA33597.1; PID:G42284

R; Iwamoto, A.; Miki, J.; Maeda, M.; Futai, M.
 J. Biol. Chem. 265, 5043-5048, 1990
 A>Title: H+-ATPase gamma-subunit of *Escherichia coli*: Role of the conserved carboxyl-ter
 A/Reference number: I55328; MUID:90202983; PMID:2138624
 A/Accession: I55328
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 261-287 <RE2>
 A/Cross-references: GB:M34095; NID:G148146; PIDN:AAA24742.1; PID:G148147
 R; Tang, C.; Wilkens, S.; Capaldi, R.A.
 J. Biol. Chem. 269, 4467-4472, 1994
 A>Title: Structure of the gamma subunit of *Escherichia coli* F-1 ATPase probed in trypsin
 A/Reference number: A53160; MUID:94140880; PMID:7508444
 A/Accession: A53160
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 2-87; 82, 204-208, 214-220 <TRAN>
 R; Blatter, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A>Title: The complete genome sequence of *Escherichia coli* K-12.
 A/Reference number: A64720; MUID:97426617; PMID:9278503
 A/Accession: F65176
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-287 <BLAT>
 A/Cross-references: GB:AE000450; GB:U00096; NID:G1790166; PIDN:AACT6756.1; PID:G1790171
 A/Experimental source: strain K-12, substrain MG1655
 C/Genetics:
 A/Gene: *atpG*; uncG
 A/Map position: 84 min
 C/Superfamily: H+-transporting ATP synthase gamma chain
 C/Keywords: *atp* biosynthesis; hydrolase; membrane-associated complex

Query Match 69.3%; Score 991; DB 1; Length 287;
 Best Local Similarity 66.1%; Pred. No. 1.1e-57;
 Matches 191; Conservative 47; Mismatches 49; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASQWRKTOERMSSSRPYSPTIRNVISHVSKAT 60
 Db 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASQWRKTOERMSSSRPYSPTIRNVISHVSKAT 60

QY 61 IYKHPFLVDREVKKVGMIVSTDRGLCGGLNVLNFKTLNEMKEWKEKDVSVQSLIGS 120
 Db 61 LEYRHPYLEERAKRQVGIITDRGLCGGLNVLNFKTLNEMKEWKEKDVSVQSLIGS 120

QY 121 KSINFPQSLGKILITQDSIGDTPSVQGLGSVNSMIDAYKGEVDVYLVYKFNFTMS 180
 Db 121 KGVSPFNSVGGVVAQVGTGMGNPSLSLIGPVKVMQAYDEGLDKLYIVSNKFNFTMS 180

QY 181 QKPVLEKLIPLPELDNDELGERKQWVDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 Db 181 QVPTISQLPLPASDDDL--KHSWDVLYEPDPKALLDILLRYYVESQVYQGVVENLAS 238

QY 241 EQARWVAMKATDNAGNLNELQVYNKARQASITNELNEIVAGAAAI 289
 Db 239 EQARWVAMKATDNAGSLIKELQVYNKARQASITQELTEIVSGAAAV 287

RESULT 4

C91213
 membrane-bound ATP synthase gamma-subunit *atpG* [imported] - *Escherichia coli* (strain O1;
 C/Species: *Escherichia coli*
 C/Date: 18-Jul-2001; #sequence, revision 18-Jul-2001; #text, change 09-Jul-2004
 C/Accession: C91213
 R; Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
 A/Reference number: A99629; MUID:21156231; PMID:11258796
 A/Accession: C91213
 A/Status: preliminary
 A/Molecule type: DNA

A:Residues: 1-287 <HAY>
A:Cross-references: UNIPROT:P00837; GB:BA000007; PIDN:BA038098.1; PID:g13364150; GSPDB:G
A:Experimental source: strain O157:H7, substrain RIMD 050952
C:Genetics:
A:Gene: ECS4675
C:Superfamily: H+-transporting ATP synthase gamma chain

Query Match 69.3%; Score 991; DB 2; Length 287;
Best Local Similarity 66.1%; Pred. No. 1.1e-57;
Matches 191; Conservative 47; Mismatches 49; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKRTQERMSSSRPSETIRNVISHVSKAT 60
DB 1 MAGAKEIRSKIASVQNTQKITKAMEMVAASQKRSQDRMAASRPYAEITMKVIGHLAHN 60

QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKVLNEMKEWKEKDVSVQLSLIGS 120
DB 61 LEYKHPYLEDRDVKRGVLYVSTDRGLCGGLNINLFKLLAEMKWTWDKGVCQDLAMIGS 120

QY 121 KSNFFQSLGILKLTQDSGIGDTPSVSEQLIGSVNSMIDAYKKGVEDVWVLYVYNKFINTWMS 180
DB 121 KGVSEFNSVGNVVAQVTGNGDPSSELLGPVKVLMQAYDEGRDKLYIVSNKFINTWMS 180

QY 181 QKPVLEKLIPELNDDELGERKQVNDYIYEPDAKVLLENLLVRYLESQVYQAAVENLAS 240
DB 181 QVPTISQLPLPASDDDL--KHKSWDYLYEPDPKALLDTLLRRYVESQVYQGVVENLAS 238

QY 241 EQARVAMKAAATDNAGLINELQLVYNKARQASITNELNEIVAGAAAI 289
DB 239 EQARVAMKAAATDGGSLIKELQLVYNKARQASITQELTEIVSGAAAV 287

RESULT 5
membrane-bound ATP synthase gamma-subunit AtpG [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D86059
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamocousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:1206551
A:Accession: D86059
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-287 <STO>
A:Cross-references: UNIPROT:P00837; GB:AE005174; NID:g12518590; PIDN:AAG58936.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: atpG
C:Superfamily: H+-transporting ATP synthase gamma chain

Query Match 69.3%; Score 991; DB 2; Length 287;
Best Local Similarity 66.1%; Pred. No. 1.1e-57;
Matches 191; Conservative 47; Mismatches 49; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKRTQERMSSSRPSETIRNVISHVSKAT 60
DB 1 MAGAKEIRSKIASVQNTQKITKAMEMVAASQKRSQDRMAASRPYAEITMKVIGHLAHN 60

QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKVLNEMKEWKEKDVSVQLSLIGS 120
DB 61 LEYKHPYLEDRDVKRGVLYVSTDRGLCGGLNINLFKLLAEMKWTWDKGVCQDLAMIGS 120

QY 121 KSNFFQSLGILKLTQDSGIGDTPSVSEQLIGSVNSMIDAYKKGVEDVWVLYVYNKFINTWMS 180
DB 121 KGVSEFNSVGNVVAQVTGMDGNLSLSLIGPVKVMQAYDEGRDKLYIVSNKFINTWMS 180

QY 181 QKPVLEKLIPELNDDELGERKQVNDYIYEPDAKVLLENLLVRYLESQVYQAAVENLAS 240
DB 181 QVPTISQLPLPASDDDL--KHKSWDYLYEPDPKALLDTLLRRYVESQVYQGVVENLAS 238

C;Superfamily: H⁺-transporting ATP synthase gamma chain

Query Match 68.4%; Score 978; DB 2; Length 287;
 Best Local Similarity 64.7%; Pred. No. 8e-57;
 Matches 187; Conservative 51; Mismatches 49; Indels 2; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKTKQERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKTKQERMSSSRPYSETIRNVISHVSKAT 60
 QY 61 IGYKHPFLVDREVKKVMIVVSTDRGLCGGLNVNLFKTVLNEMKEKQDVSVQLSLIGS 120
 DB 61 LEYKHPLEERAKVGVYIVVSTDRGLCGGLNVNLFKTVLNEMKEKQDVSVQLSLIGS 120
 QY 121 KSNFFQSLGKILTPQDSIGDTPSVEQLIGSVNSMIDAYKKGVEDVYVLYNKFINTMS 180
 DB 121 KGVSVFNSVGGNVVAVQVTGMDGNFSLSELIGPVKVMQLQAYDEGLDKLYVSNKFINTMS 180
 QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QVPTIQQLPLPASEDDLL--KRTANDYIYEPDPKALLDTLLRRYVESQVYQGVVENLAS 238
 QY 241 EQAARMVAMKAATDNAGNLINELQVLYNKAQASITNELNEIVAGAAAI 289
 DB 239 EQAARMVAMKAATDNAGSLIKELQVLYNKAQASITQELTEIVSGAAAV 287

RESULT 8

S06081
 H⁺-transporting two-sector ATPase (EC 3.6.3.14) gamma chain - Vibrio alginolyticus
 C;Species: Vibrio alginolyticus
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: S06081
 R;Krumholz, L.R.; Esser, U.; Simoni, R.D.
 Nucleic Acids Res. 17, 7993-7994, 1989
 A;Title: Nucleotide sequence of the unc operon of Vibrio alginolyticus.
 A;Reference number: S06075; MUID:90016889; PMID:2529481
 A;Accession: S06081
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-288 <RRU>
 A;Cross-references: UNIPROT:P12990; EMBL:X16050; NID:G48331; PID:G48338
 C;Genetics:
 A;Gene: uncG
 C;Superfamily: H⁺-transporting ATP synthase gamma chain
 C;Keywords: ATP biosynthesis; hydrolase; membrane-associated complex

Query Match 68.3%; Score 976.5; DB 1; Length 288;
 Best Local Similarity 63.3%; Pred. No. 1e-56;
 Matches 183; Conservative 56; Mismatches 43; Indels 1; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKTKQERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKTKQERMSSSRPYSETIRNVISHVSKAT 60
 QY 61 IGYKHPFLVDREVKKVMIVVSTDRGLCGGLNVNLFKTVLNEMKEKQDVSVQLSLIGS 120
 DB 61 LEYRHPYLEERAKVGVYIVVSTDRGLCGGLNVNLFKTVLNEMKEKQDVSVQLSLIGS 120
 QY 121 KSNFFQSLGKILTPQDSIGDTPSVEQLIGSVNSMIDAYKKGVEDVYVLYNKFINTMS 180
 DB 121 KATAPFKHGGAAVAAQVSGLGNSFLELDIGSVGWLKKYDEGBELDLYVFNKFNVTWV 180
 QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QPTDQQLPLPKSDSEM-QREHSWDYIYEPKPLDITLLRRYVESQVYQGVVENLAS 239
 QY 241 EQAARMVAMKAATDNAGNLINELQVLYNKAQASITNELNEIVAGAAAI 289
 DB 240 EQAARMVAMKAATDNATNLIDDLVLYNKAQASITQELSELVEIVGGAAV 288

RESULT 9

1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKTKQERMSSSRPYSETIRNVISHVSKAT 60

D82952
 ATP synthase gamma chain PA5555 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C;Accession: D82952
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim.
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: D82952
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-286 <STO>
 A;Cross-references: UNIPROT:Q9H119; GB:AE004967; GB:AE004091; NID:G9951884; PIDN:AAG069.
 A;Experimental source: strain PAO1
 C;Genetics:
 A;Gene: atpG; PA5555
 C;Superfamily: H⁺-transporting ATP synthase gamma chain

Query Match 64.8%; Score 926.5; DB 2; Length 286;
 Best Local Similarity 62.3%; Pred. No. 1.8e-53;
 Matches 180; Conservative 50; Mismatches 56; Indels 3; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKTKQERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKTKQERMSSSRPYSETIRNVISHVSKAT 60
 QY 61 IGYKHPFLVDREVKKVMIVVSTDRGLCGGLNVNLFKTVLNEMKEKQDVSVQLSLIGS 120
 DB 61 PEYRPFVWEVREVKVGVYIVVSSDRGLCGGLNVNLFKTVLNEMKEKQDVSVQLSLIGS 120
 QY 121 KSNFFQSLGKILTPQDSIGDTPSVEQLIGSVNSMIDAYKKGVEDVYVLYNKFINTMS 180
 DB 121 KGASPFSGGNGVAAISHLGEPSINDLIGSVKVMQLDAYLEGRIDELFVSNKFNVTMT 180
 QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QKPTVEQLPLVADDDQL---KHWDYLYEPDAKSLDGLLVRYVESQVYQAAVENNAC 237
 QY 241 EQAARMVAMKAATDNAGNLINELQVLYNKAQASITNELNEIVAGAAAI 289
 DB 238 EQAARMVAMKAATDNAGELISDLQVLYNKAQASITQELISEIVGGAAAV 286

RESULT 10

G84930
 H⁺-transporting two-sector ATPase (EC 3.6.3.14) gamma chain [imported] - Buchnera sp. (ε
 C;Species: Buchnera sp.
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 03-Jun-2002
 C;Accession: G84930
 R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. f
 A;Reference number: A84930; MUID:20445173; PMID:10993077
 A;Accession: G84930
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-290 <STO>
 A;Cross-references: GB:AP000398; GSPDB:GN00144
 A;Experimental source: strain APS
 C;Genetics:
 A;Gene: atpG; BU007
 C;Superfamily: H⁺-transporting ATP synthase gamma chain
 C;Keywords: hydrolase

Query Match 56.5%; Score 808; DB 2; Length 290;
 Best Local Similarity 52.8%; Pred. No. 1e-45;
 Matches 152; Conservative 60; Mismatches 75; Indels 2; Gaps 1;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 09:47:28 ; Search time 193 Seconds
(without alignments)
861.571 Million cell updates/sec

Title: US-09-545-199F-4

Perfect score: 1429

Sequence: 1 MAGAKERTKIASVKSTQKI.....ARQASITNELNIVAGAAAI 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1429	100.0	289	1	ATPG_PASMU	Q916b6 pasteurella
2	1173	82.1	289	1	ATPG_HAEIN	P43716 haemophilus
3	1121.5	78.5	288	2	Q7VPP1	Q7vpp1 haemophilus
4	1004	70.3	287	2	Q7NA93	Q7na93 photorhabdu
5	998.5	69.9	288	2	Q9KNH4	Q9knh4 vibrio chol
6	993	69.5	287	2	Q6CVJ4	Q6cvj4 erwiaia car
7	991	69.3	287	1	ATPG_ECOLI	P00837 escherichia
8	991	69.3	287	2	Q9PFI4	Q9pfi4 salmonella
9	989	69.2	287	2	Q8Z9S5	Q8z9s5 yersinia pe
10	989	69.2	287	2	AA564168	AA564168 yersinia
11	988	69.1	287	2	Q8ZK98	Q8zk98 salmonella
12	982.5	68.8	286	2	Q8E8B9	Q8e8b9 shewanella
13	980.5	68.6	288	2	Q7MGH9	Q7mgh9 vibrio vuln
14	980.5	68.6	288	2	Q8DDG9	Q8ddg9 vibrio vuln
15	978	68.4	287	2	Q8Z2O5	Q8z2o5 salmonella
16	976.5	68.3	288	1	ATPG_VIBAL	P12990 vibrio algi
17	970.5	67.9	288	2	Q87KA7	Q87ka7 vibrio para
18	947.5	66.3	288	2	Q6LLG7	Q6llg7 photobacter
19	947.5	66.3	288	2	CAG21861	Cag21861 photobact
20	931.5	65.2	287	2	Q8VW78	Q8vw78 colwellia m
21	926.5	64.8	286	2	Q9HT19	Q9ht19 pseudomonas
22	915	64.0	291	2	Q5LXZ7	Q5lxz7 photobacter
23	915	64.0	291	2	CAG22008	Cag22008 photobact
24	896.5	62.7	286	2	Q88BX3	Q88bx3 pseudomonas
25	889	62.2	289	2	Q6FFK1	Q6ffk1 acinetobact
26	885.5	62.0	286	2	Q87TT3	Q87tt3 pseudomonas
27	885.5	62.0	294	2	Q82XP9	Q82xp9 nitrosomona
28	885	61.9	291	2	Q8XU75	Q8xu75 ralstonia s
29	847.5	59.3	289	2	Q83AF5	Q83af5 coxiella bu
30	838	58.6	287	2	Q8FC26	Q8fc26 xanthomonas
31	830	58.1	287	2	Q8P6G6	Q8p6g6 xanthomonas

32	808	56.5	290	1	ATPG_BUCAI	P57123 buchnera ap
33	805	56.3	301	2	Q7VU45	Q7vu45 bordetella
34	804	56.3	301	2	Q7WEM8	Q7wem8 bordetella
35	802	56.1	301	2	Q7WJA9	Q7wj99 thiodetella
36	799.5	55.9	298	1	ATPG_THIFE	P41169 thiodetella
37	795	55.6	287	2	Q9PE84	Q9pe84 xylobacillu
38	791	55.4	287	2	Q87E89	Q87e89 xylobacillu
39	790.5	55.3	291	1	ATPG_BUCAP	Q87e89 xylobacillu
40	783.5	54.8	288	2	Q7VU77	Q7vu77 candidatus
41	754	52.8	291	2	Q9NW71	Q9nw71 neisseria m
42	753	52.7	291	2	Q9UXQ1	Q9uxq1 neisseria m
43	751.5	52.6	275	2	Q7P096	Q7p096 chromobacte
44	747	52.3	290	2	Q8RQ80	Q8rq80 buchnera ap
45	693.5	48.5	287	2	Q8D3J4	Q8dj34 wiggleswort

ALIGNMENTS

RESULT 1

ATPG_PASMU	ID	ATPG_PASMU	STANDARD;	PRT;	289 AA.
AC	Q9L6B6;				
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	05-JUL-2004	(Rel. 44, Last annotation update)			
DE	ATP synthase gamma chain (EC 3.6.3.14).				
GN	Name=atpg; OrderedLocusNames=PM1493;				
OS	Pasteurella multocida.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;				
OC	Pasteurellaceae; Pasteurella.				
OX	NCBI_TaxID=747;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Fuller T.E., Kennedy M.J., Lowery D.E.;				
RT	"Identification of Pasteurella multocida virulence genes in a				
RT	septicemic mouse model using signature-tagged mutagenesis."				
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Pm70;				
RX	MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;				
RA	May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;				
RT	"Complete genomic sequence of Pasteurella multocida Pm70."				
Proc.	Natl Acad Sci. U.S.A. 98:3460-3465(2001).				
CC	!- FUNCTION: Produces ATP from ADP in the presence of a proton				
CC	gradient across the membrane. The gamma chain is believed to be				
CC	important in regulating ATPase activity and the flow of protons				
CC	through the CF(0) complex.				
CC	!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +				
CC	H(+) (Out).				
CC	!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic				
CC	core - and CF(0) - the membrane proton channel. CF(1) has five				
CC	subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)				
CC	has three main subunits: a, b and c.				
CC	!- SIMILARITY: Belongs to the ATPase gamma chain family.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL; AF237922; AA968408.1; -				
DR	EMBL; AF006186; AA003577.1; -				
DR	HSFP; P00837; IFS0.				
DR	InterPro; IPR000131; ATPase_gamma.				
DR	Pfam; PF00231; ATP-synt; 1.				
DR	PRINTS; PR00126; ATPASEGAMMA.				
DR	TIGRFAMs; TIGR01146; ATPsyn_F1gamma; 1.				
DR	PROSITE; PS00153; ATPASE_GAMMA; 1.				

KW ATP synthesis; CF(1); Complete proteome; Hydrogen ion transport;
 KW Hydrolyase.
 SQ SEQUENCE 289 AA; 32095 MW; 1E18862B4EA9F70 CRC64;
 Query Match 100.0%; Score 1429; DB 1; Length 289;
 Best Local Similarity 100.0%; Pred. No. 8.6e-86;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASKRKTOERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASKRKTOERMSSSRPYSETIRNVISHVSKAT 60
 QY 61 IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTVLNEMKEWKEKDVSVQLSLIGS 120
 DB 61 IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTVLNEMKEWKEKDVSVQLSLIGS 120
 QY 121 KSNFFQSLGKILTDQSGIGDTPSPVEQLIGSVNSMIDAYKGEVDVYLVNKFINTMS 180
 DB 121 KSNFFQSLGKILTDQSGIGDTPSPVEQLIGSVNSMIDAYKGEVDVYLVNKFINTMS 180
 QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 QY 241 EQAARVMVAKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 289
 DB 241 EQAARVMVAKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 289

RESULT 2
 ATPG_HABIN STANDARD; PRT; 289 AA.
 AC P43776;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE ATP synthase gamma chain (EC 3.6.3.14).
 GN Name=atpG; OrderedLocusNames=H10480;
 OS Haemophilus influenzae
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kierlavage A.R., Ehalt C.O., Tomb J.-P., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Saudek D.M., Brandon R.C.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RL Science 269:496-512(1995).
 RC -!- FUNCTION: Produces ATP from ADP in the presence of a proton
 CC gradient across the membrane. The gamma chain is believed to be
 CC important in regulating ATPase activity and the flow of protons
 CC through the CF(0) complex.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
 CC core - and CF(0) - the membrane proton channel. CF(1) has five
 CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
 CC has three main subunits: a, b and c.
 CC -!- SIMILARITY: Belongs to the ATPase gamma chain family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 DR EMBL; U32730; AAC22138.1; -.
 DR F1R; E64071; E64071.
 DR HSSP; P00837; 1FSO.
 DR TIGR; H10480; -.
 DR InterPro; IPR000131; ATPase_gamma.
 DR Pfam; PF00231; ATP-synt; 1.
 DR PRINTS; PR00126; ATPASEGAMMA.
 DR TIGRFAMs; TIGR01146; ATPsyn_Figamma; 1.
 DR PROSITE; PS00153; ATPASE_GAMMA; 1.
 KW ATP synthesis; CF(1); Complete proteome; Hydrogen ion transport;
 KW Hydrolyase.
 SQ SEQUENCE 289 AA; 32069 MW; 622CEA682F37FD00 CRC64;
 Query Match 82.1%; Score 1173; DB 1; Length 289;
 Best Local Similarity 76.5%; Pred. No. 5.1e-69;
 Matches 221; Conservative 42; Mismatches 26; Indels 0; Gaps 0;

QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASKRKTOERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASKRKTOERMSSSRPYSETIRNVISHVSKAT 60
 QY 61 IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTVLNEMKEWKEKDVSVQLSLIGS 120
 DB 61 IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTVLNEMKEWKEKDVSVQLSLIGS 120
 QY 121 KSNFFQSLGKILTDQSGIGDTPSPVEQLIGSVNSMIDAYKGEVDVYLVNKFINTMS 180
 DB 121 KSNFFQSLGKILTDQSGIGDTPSPVEQLIGSVNSMIDAYKGEVDVYLVNKFINTMS 180
 QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QKPVVQVLPLPEKDDHLNERQQTWDYLYEPKPKVLLDLSLLVRYLESQVYQAAVENLAS 240
 QY 241 EQAARVMVAKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 289
 DB 241 EQAARVMVAKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 289

RESULT 3
 Q7VPP1 PRELIMINARY; PRT; 288 AA.
 AC Q7VPP1;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE ATP synthase gamma chain.
 GN Name=atpG; OrderedLocusNames=HD0009;
 OS Haemophilus ducreyi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=730;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=35000HP / ATCC 700724;
 RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
 RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
 RT "The complete genome sequence of Haemophilus ducreyi.";
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017151; AAP95033.1; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
 DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
 DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.
 DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
 DR InterPro; IPR000131; ATPase_gamma.
 DR Pfam; PF00231; ATP-synt; 1.
 DR TIGRFAMs; TIGR01146; ATPsyn_Figamma; 1.
 DR PROSITE; PS00153; ATPASE_GAMMA; 1.

KW Complete proteome.
SQ SEQUENCE 288 AA; 31953 MW; 6CC7342402D4C977 CRC64;

Query Match 78.5%; Score 1121.5; DB 2; Length 288;
Best Local Similarity 73.4%; Pred. No. 1.2e-65;
Matches 212; Conservative 48; Mismatches 28; Indels 1; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKAMENVAASVKRKTQERMSSSRPYSETIRNVISHVSKAT 60
DB 1 MAGAKEIRTKIASVKSTOKITKAMENVAASVKRKTQERMSSSRPYSETIRNVISHVSKAT 60

QY 61 IGYKHPFLVDREYKVKGMIVSTDRGLCGGLNVLNFKTVLNEMKEWKDKVSVLSLIGS 120
DB 61 IGYKHPFLVDREYKVKGMIVSTDRGLCGGLNVLNFKTVLNEMKEWKDKVSVLSLIGS 120

QY 121 KSNFQSLGKILTDGSGIGDTPSVEQLIGSVNSMIDAYKKGEVDVYLVYKFNITMS 180
DB 121 KSNFQSLGKILTDGSGIGDTPSVEQLIGSVNSMIDAYKKGEVDVYLVYKFNITMS 180

QY 181 KQVLEKLIPLPELDNDELGERKQWVDIYEPDQVLLDNLVRYLESQVYQAAVENLAS 240
DB 181 KQVLEKLIPLPELDNDELGERKQWVDIYEPDQVLLDNLVRYLESQVYQAAVENLAS 240

QY 241 EQAARWVAKAATDNAGNINELQVYNKAROASITNENIIVAGAAAI 289
DB 241 EQAARWVAKAATDNAGNINELQVYNKAROASITNENIIVAGAAAI 289

QY 240 EQAARWVAKAATDNAGNINELQVYNKAROASITNENIIVAGAAAI 288
DB 240 EQAARWVAKAATDNAGNINELQVYNKAROASITNENIIVAGAAAI 288

RESULT 4
Q7NA93
ID Q7NA93 PRELIMINARY; PRT; 287 AA.
AC Q7NA93;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE ATP synthase gamma chain.
GN Names-atpG; OrderedLocusNames=plu0041;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22557627; PubMed=14528314;
RA Duchaud E., Rusnok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taurit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freysinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Siguier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.,
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens."
RL Nat. Biotechnol. 21:1307-1313(2003).
DR EMBL; BX571859; CAE12336.1; -.
DR PhotoList; plu0041; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.
DR GO; GO:0045986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro; IPR000131; ATPase_gamma.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; PR00126; ATPSEGAMMA.
DR TIGRfams; TIGR01146; ATPsyn_Figamma; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 31644 MW; D957DBEF07B7F415 CRC64;

Query Match 70.3%; Score 1004; DB 2; Length 287;
Best Local Similarity 67.5%; Pred. No. 6.1e-58;
Matches 195; Conservative 48; Mismatches 44; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKAMENVAASVKRKTQERMSSSRPYSETIRNVISHVSKAT 60
DB 1 MAGAKEIRTKIASVKSTOKITKAMENVAASVKRKTQERMSSSRPYSETIRNVISHVSKAT 60

QY 61 IGYKHPFLVDREYKVKGMIVSTDRGLCGGLNVLNFKTVLNEMKEWKDKVSVLSLIGS 120
DB 61 IGYKHPFLVDREYKVKGMIVSTDRGLCGGLNVLNFKTVLNEMKEWKDKVSVLSLIGS 120

QY 121 KSNFQSLGKILTDGSGIGDTPSVEQLIGSVNSMIDAYKKGEVDVYLVYKFNITMS 180
DB 121 KSNFQSLGKILTDGSGIGDTPSVEQLIGSVNSMIDAYKKGEVDVYLVYKFNITMS 180

QY 181 KQVLEKLIPLPELDNDELGERKQWVDIYEPDQVLLDNLVRYLESQVYQAAVENLAS 240
DB 181 KQVLEKLIPLPELDNDELGERKQWVDIYEPDQVLLDNLVRYLESQVYQAAVENLAS 240

QY 241 EQAARWVAKAATDNAGNINELQVYNKAROASITNENIIVAGAAAI 289
DB 241 EQAARWVAKAATDNAGNINELQVYNKAROASITNENIIVAGAAAI 289

QY 239 EQAARWVAKAATDNAGNINELQVYNKAROASITNENIIVAGAAAI 287
DB 239 EQAARWVAKAATDNAGNINELQVYNKAROASITNENIIVAGAAAI 287

RESULT 5
Q9KNH4
ID Q9KNH4 PRELIMINARY; PRT; 288 AA.
AC Q9KNH4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE ATP synthase F1, gamma subunit.
GN OrderedLocusNames=VC2765;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey R.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AB004342; AAF95904.1; -.
DR PIR; G82036; G82036.
DR HSSP; P00837; 1FS0.
DR TIGR; VC2765; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro; IPR000131; ATPase_gamma.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; PR00126; ATPSEGAMMA.
DR TIGRfams; TIGR01146; ATPsyn_Figamma; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; 1.
KW Complete proteome.
SQ SEQUENCE 288 AA; 31862 MW; 2CCE2F03B9982DBF CRC64;

Query Match 69.9%; Score 998.5; DB 2; Length 288;
Best Local Similarity 65.1%; Pred. No. 1.4e-57;
Matches 188; Conservative 54; Mismatches 46; Indels 1; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKAMENVAASVKRKTQERMSSSRPYSETIRNVISHVSKAT 60
DB 1 MAGAKEIRTKIASVKSTOKITKAMENVAASVKRKTQERMSSSRPYSETIRNVISHVSKAT 60

P00837; P00838;
21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
01-OCT-2004 (Rel. 45, Last annotation update)
ATP synthase gamma chain (EC 3.6.3.14).
Name=atpG; Synonym=uncG, pApC
OrderedLocNames=B3733, c4659, #5231, EC#4675, SF3813, S3955;
Escherichia coli, pApC
Escherichia coli O6,
Escherichia coli C157:H7, and
Shigella flexneri.
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
NCBI_taxID=562, 217992, 83334, 623;
[1]
SEQUENCE FROM N.A.
SPECIES=E.coli;
MEDLINE=8511806; PubMed=6395859;
Walker J.E., Gay N.J., Sarate M., Eberle A.N.;
"DNA sequence around the Escherichia coli unc operon. Completion of
the sequence of a 17 kilobase segment containing asnA, orfC, unc, glms
and phoS";
Biochem. J. 224:799-815(1984).
[2]
SEQUENCE FROM N.A.
SPECIES=E.coli;
SPECIES=E.coli;
MEDLINE=82059507; PubMed=6272217;
Sarate M., Gay N.J., Eberle A., Runswick M.J., Walker J.E.;
"The atp operon: nucleotide sequence of the genes for the gamma, beta,
and epsilon subunits of Escherichia coli ATP synthase.";
Nucleic Acids Res. 9:5287-5296(1981).
[3]
SEQUENCE FROM N.A.
SPECIES=E.coli;
SPECIES=E.coli;
MEDLINE=82134798; PubMed=6277310;
Kanazawa H., Kayano T., Mabuchi K., Futai M.;
"Nucleotide sequence of the genes coding for alpha, beta and gamma
subunits of the proton-translocating ATPase of Escherichia coli.";
Biochem. Biophys. Res. Commun. 103:604-612(1981).
[4]
SEQUENCE FROM N.A.
SPECIES=E.coli;
SPECIES=E.coli;
MEDLINE=83176724; PubMed=6301339;
Kanazawa H., Futai M.;
"Structure and function of H+-ATPase: what we have learned from
Escherichia coli H+-ATPase.";
Ann. N. Y. Acad. Sci. 402:45-64(1982).
[5]
SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
MEDLINE=93351543; PubMed=7686882;
Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
"DNA sequence and analysis of 136 kilobases of the Escherichia coli
genome: organizational symmetry around the origin of replication.";
Genomics 16:551-561(1993).
[6]
SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
MEDLINE=42398234; PubMed=12471157; DOI=10.1073/pnas.252529799;
Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
Raab D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
"Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
[7]
SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
MEDLINE=21074335; PubMed=11206551; DOI=10.1038/35054089;
Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klein L., Borczyk A.,
Fleishman J.A., Goulding M.P., White O., Salzberg S.L.,
Karp P., Fraser H., Blattner F.R.;
"Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and comparison with other serotypes.";
Science 284:1907-1911(1999).

SPECIES: *E.COLI*; STRAIN: O157:H7 / EDL933 / ATCC 700927 / EHCC;
MEDLINE: 12017435; PubMed: 11206551; DOI: 10.1038/35054089;
Perrin N.A., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouisis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
RL Nature 409:529-533(2001).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=*E.coli*; STRAIN=O157:H7 / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki A.C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [9]
RP SEQUENCE OF 261-287 FROM N.A.
RC SPECIES=*E.coli*;
RX MEDLINE=90202983; PubMed=2138624;
RA Iwamoto A., Miki J., Maeda M., Futai M.;
RT "H(+) ATPase gamma subunit of *Escherichia coli*. Role of the conserved
RT carboxyl-terminal region.";
RL J. Biol. Chem. 265:5043-5048(1990).
RN [10]
RP SEQUENCE OF 2-9; 72-81; 203-208 AND 214-220.
RC SPECIES=*E.coli*;
RX PubMed=7508444;
RA Tang C., Wilkens S., Capaldi R.A.;
RT "Structure of the gamma subunit of *Escherichia coli* F1 ATPase probed
RT in trypsin digestion and biotin-avidin binding studies.";
RL J. Biol. Chem. 269:4467-4472(1994).
RN [11]
RP SEQUENCE FROM N.A.
RC SPECIES=*S.flexneri*; STRAIN=301 / Serotype 2a;
RX MEDLINE=2227406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
RT through comparison with genomes of *Escherichia coli* K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [12]
RP SEQUENCE FROM N.A.
RC SPECIES=*S.flexneri*; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of *Shigella*
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (4.4 ANGSTROMS).
RC SPECIES=*E.coli*;
RX MEDLINE=20040613; PubMed=10570135;
RA Hausarth A.C., Gruber G., Matthews B.W., Capaldi R.A.;
RT "Structural features of the gamma subunit of the *Escherichia coli* F(1)
RT ATPase revealed by a 4.4-A resolution map obtained by X-ray
RT crystallography.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:13697-13702(1999).
CC -!- FUNCTION: Produces ATP from ADP in the presence of a proton
CC gradient across the membrane. The gamma chain is believed to be
CC important in regulating ATPase activity and the flow of protons
CC through the CF(0) complex.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC

CC has three main subunits: a, b and c.
CC -!- SIMILARITY: Belongs to the ATPase gamma chain family.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC and various other errors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; X01631; CAA25781.1; -
DR EMBL; J01594; AAA24736.1; ALT_FRAME.
DR EMBL; V00367; CAA23526.1; -
DR EMBL; V00312; CAA23597.1; ALT_FRAME.
DR EMBL; M25464; AAN83874.1; -
DR EMBL; L10328; AAG62085.1; -
DR EMBL; AE000450; AAC76756.1; -
DR EMBL; AE016769; AAN83091.1; -
DR EMBL; AE005605; AAG58936.1; -
DR EMBL; AP002566; BAB38098.1; -
DR EMBL; M34095; AAA24742.1; -
DR EMBL; AE015388; AAN45353.1; -
DR EMBL; AE016991; AAP18944.1; -
DR PIR; A01038; PWECC.
DR PIR; C91213; C91213.
DR PIR; D86059; D86059.
DR PDB; 1D8S; X-ray; G=-.
DR PDB; 1F80; X-ray; G=19-248.
DR EcoGene; EB0102; -
DR EcoGene; EG10104; atpG.
DR InterPro; IPR000131; ATPase_gamma.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; PR00126; ATPASEGAMMA.
DR TIGRFAMS; TIGR01146; ATPsyn_Flgamma; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; 1.
DR 3D-structure; ATP synthase; CF(1); Complete proteome;
KW Direct protein sequencing; Hydrogen ion transport; Hydrolase.
FT HELIX 20 57
FT HELIX 66 68
FT STRAND 75 81
FT TURN 89 90
FT TURN 91 108
FT TURN 109 110
FT STRAND 112 118
FT HELIX 120 129
FT STRAND 133 137
FT TURN 141 142
FT HELIX 147 161
FT TURN 162 163
FT STRAND 168 177

Query Match 69.3%; Score 991; DB 1; Length 287;
Best Local Similarity 66.1%; Pred. No. 4.3e-57;
Matches 191; Conservative 47; Mismatches 49; Indels 2; Gaps 1;
QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASVKTKQTERMSRRPYSETIRNIVSHVSKAT 60
Db 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASVKTKQTERMSRRPYSETIRNIVSHVSKAT 60
QY 61 IGYKHPLVDREVKKYGMIVVSTDRGLCGGLNVLNFKTVLNMKWKEDKVSQVLSLGS 120
Db 61 LEYKHPLYEDRVKRVGLVSVSDRGLCGGLNVLNFKTVLNMKWKEDKVSQVLSLGS 120
QY 121 KSNFFPSGLKILTDQSGIDGTPSVEQLIGSVNGMIDAYKKGEVDVYLVYKNEFTMS 180
Db 121 KGVSFNFVSGNVAVQVTGMDGNFSLSELIGPVKVLQAYDEGRDLKLYIVKNEFTMS 180
QY 181 QKPVLEKLIPLPELDNDELGERKQVNDVIYEDAKVLLDNLVLYLESQVTOAAVENLAS 240
Db 181 QVPTISQLPLPASDDDL--KHKSWDLYLEPDKALLDTLLRRYVESQVYGVVENLAS 238

```

QY 241 EQARWVAMKAAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
DB 239 EQARWVAMKAAATDNGSLIKELQVYNKARQASITQELTEIVSGAAV 287

RESULT 8
Q9RFL4
ID Q9RFL4 PRELIMINARY; PRT; 287 AA.
AC Q9RFL4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ATP synthase subunit gamma.
GN Name=atpG;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TA98;
RA Kim H.-K., Heo N.-J., Ghim S.-Y., Song B.-H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188265; AA19361.1; -.
DR HSP; P00837; 1FS0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity. .; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. .; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro; IPR000131; ATPase gamma.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; P00126; ATPSEGAMMA.
DR TIGRfam; TIGR01146; ATPsyn Figamma; 1.
DR PROSITE; PS00153; ATPASE GAMMA; 1.
DR PROSITE; PS00041; HTH ARAC FAMILY_1; UNKNOWN_1.
KW Complete proteome, Hydrolase.
SQ SEQUENCE 287 AA; 31578 MW; A0B3E8D41B3B80F6 CRC64;

Query Match 69.3%; Score 991; DB 2; Length 287;
Best Local Similarity 66.1%; Pred. No. 4.3e-57;
Matches 191; Conservative 47; Mismatches 49; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKANEMVAASQKRTQERMSSRPYSYETTRNVISHVSKAT 60
DB 1 MAGAKEIRSKIASVQNTQKITKANEMVAASQKRSQDRMAASRPYAEYTKRWVIGHLAGN 60
QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNEMKEKEDSVQSLIGS 120
DB 61 LEYKHPYLEDREVKKVGLVYVSTDRGLCGGLNVNLFKLLAEMKKTWDRKGVQCDLAWIGS 120
QY 121 KSNFFOSLGIKILTQDSIGDTPSVQQLIGSVNSMIDAYKKGEVDVVLVYNKFTINTMS 180
DB 121 KGVSEFFDSVGVNVAQVTGMDNPSSELLGPVKVWLQAYDEGRDLKLYIVSNKFTINTMS 180
QY 161 QKPVLEKLIPLPDLNDELGERQVWDYIYEPDAKVLNLDNLVRYLESQVYQAAVENLAS 240
DB 161 QVPTISQLLPASDDDL--KHKSMDYIYEPDPKALLDTLLRRYVESQVYQGVVENLAS 238
QY 241 EQARWVAMKAAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
DB 239 EQARWVAMKAAATDNGSLIKELQVYNKARQASITQELTEIVSGAAV 287

RESULT 9
Q8Z9S5
ID Q8Z9S5 PRELIMINARY; PRT; 287 AA.
AC Q8Z9S5; Q74PA0; Q7CFM7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE ATP synthase gamma subunit protein (EC 3.6.1.34) (Membrane-bound ATP synthase, F1 sector, gamma-subunit).
GN Name=atpG; Synonyms=papC, uncG;

```

```

GN OrderedLocusNames=YP4029, YP4122, Y4136;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titchell R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.,
RT Genome sequence of Yersinia pestis, the causative agent of plague.;
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Zhubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT Genome sequence of Yersinia pestis KIM.;
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen P., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ414160; CAC93571.1; -.
DR EMBL; AE014015; AAM87678.1; -.
DR EMBL; AE017142; AAS64168.1; -.
DR FIR; AF0500; AF0500.
DR HSP; P00837; 1FS0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity. .; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. .; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro; IPR000131; ATPase gamma.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; P00126; ATPSEGAMMA.
DR TIGRfam; TIGR01146; ATPsyn Figamma; 1.
DR PROSITE; PS00153; ATPASE GAMMA; 1.
DR PROSITE; PS00041; HTH ARAC FAMILY_1; UNKNOWN_1.
KW Complete proteome, Hydrolase.
SQ SEQUENCE 287 AA; 31577 MW; 2F124E8B7CE031CA CRC64;

Query Match 69.2%; Score 989; DB 2; Length 287;
Best Local Similarity 67.1%; Pred. No. 5.8e-57;
Matches 194; Conservative 46; Mismatches 47; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTOKITKANEMVAASQKRTQERMSSRPYSYETTRNVISHVSKAT 60
DB 1 MAGAKEIRSKIASVQNTQKITKANEMVAASQKRSQDRMAASRPYAEYTKRWVIGHLAGN 60
QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNEMKEKEDSVQSLIGS 120
DB 61 LEYKHPYLEDREVKKVGLVYVSTDRGLCGGLNVNLFKLLAEMKKTWDRKGVQCDLAWIGS 120
QY 121 KSNFFOSLGIKILTQDSIGDTPSVQQLIGSVNSMIDAYKKGEVDVVLVYNKFTINTMS 180
DB 121 KAASFFGSGVKIQAQVTGMDNPSSELLGPVKVWLQAYDEGRDLKLYIVNNKFTINTMS 180

```

```

QY 181 QKPVLEKLIPLPELDNDELGERKQWDYIYEPDAKVLNLIYRVLESQVYQAAVENLAS 240
DB 181 QEPRIWQLPLPAEDGEL--KKKSDYLYEYDPKALLDTLLRRVVSQVYQGVVENLAS 238

QY 241 EQAARWVAKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
DB 239 EQAARWVAKAATDNGSLIKELQVYNKARQASITQELTEIVGASAV 287

RESULT 10
AAS64168
ID AAS64168 PRELIMINARY; PRT; 287 AA.
AC AAS64168
DT 24-MAR-2004 (TRENBLrel. 27, Created)
DT 24-MAR-2004 (TRENBLrel. 27, Last sequence update)
DE ATP synthase gamma subunit protein.
GN ATPG OR YP4029.
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCB1_TaxID=632;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017142; AAS64168.1; -.
SQ SEQUENCE 287 AA; 31577 MW; 2F124E8B7CE031CA CRC64;

Query Match 69.2%; Score 989; DB 2; Length 287;
Best Local Similarity 67.1%; Pred. No. 5.8e-57;
Matches 194; Conservative 46; Mismatches 47; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKMKTKOERSSSRPYSETIRNVISHVSKAT 60
DB 1 MAGAKEIRSKIASVQNTQKITKAMEMVAASKMKTKOERVAASRPYAEIRVSHVHLANGN 60

QY 61 IGYKHPFLVDREVKKYGMIVVSTDRGLCGGLNVLNFKTVLNEKWKEDKVSQVLSLIGS 120
DB 61 LEYKHPYLEERDKRVKGYLVVSTDRGLCGGLNVLNFKLLAEMKWKSEKGVCECDLALIGS 120

QY 121 KSINFFQSLGKILTDQSGIGDTPSPVEQLIGSVNSMIDAYKKGEVDVYLVYVYKPFINTMS 180
DB 121 KAASFFGSGVGGKIVAVTGNQDNPSLSELIQPKVNLQAYDEGRUDKLYIYVYKPFINTMS 180

QY 181 QKPVLEKLIPLPELDNDELGERKQWDYIYEPDAKVLNLIYRVLESQVYQAAVENLAS 240
DB 181 QEPRIWQLPLPAEDGEL--KKKSDYLYEYDPKALLDTLLRRVVSQVYQGVVENLAS 238

QY 241 EQAARWVAKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
DB 239 EQAARWVAKAATDNGSLIKELQVYNKARQASITQELTEIVGASAV 287

RESULT 11
Q82KW8
ID Q82KW8 PRELIMINARY; PRT; 287 AA.
AC Q82KW8;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DE Membrane-bound ATP synthase, F1 sector, gamma-subunit.
GN Name=atpG; OrderedLocusNames=STM3866;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCB1_TaxID=602;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:853-856(2001).
DR EMBL; AE008880; AAL22724.1; -.
DR HSSP; P00837; IFSO.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro; IPR000131; ATPase_gamma.
DR Pfam; PF00231; ATP-synt; I.
DR PRINTS; PR00126; ATPASEGAMMA.
DR TIGRFBMS; TIGR01146; ATPsyn_Figamma; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; 1.
DR Complete proteome.
KW SEQUENCE 287 AA; 31555 MW; C33B55F0F5B4FA29 CRC64;

Query Match 69.1%; Score 988; DB 2; Length 287;
Best Local Similarity 65.1%; Pred. No. 6.8e-57;
Matches 188; Conservative 52; Mismatches 47; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKMKTKOERSSSRPYSETIRNVISHVSKAT 60
DB 1 MAGAKEIRSKIASVQNTQKITKAMEMVAASKMKTKOERVAASRPYAEIRVSHVHLANGN 60

QY 61 IGYKHPFLVDREVKKYGMIVVSTDRGLCGGLNVLNFKTVLNEKWKEDKVSQVLSLIGS 120
DB 61 LEYKHPYLEERDKRVKGYLVVSTDRGLCGGLNVLNFKLLADMKAWSKGVQCELAMIGS 120

QY 121 KSINFFQSLGKILTDQSGIGDTPSPVEQLIGSVNSMIDAYKKGEVDVYLVYVYKPFINTMS 180
DB 121 KGVSFFNSVGGVVAQVTVGNQDNPSLSELIQPKVNLQAYDEGRUDKLYIYVYKPFINTMS 180

QY 181 QKPVLEKLIPLPELDNDELGERKQWDYIYEPDAKVLNLIYRVLESQVYQAAVENLAS 240
DB 181 QVPTITQLPLPASEDDDL--KKKAWDYLYEYDPKALLDTLLRRVVSQVYQGVVENLAS 238

QY 241 EQAARWVAKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
DB 239 EQAARWVAKAATDNGSLIKELQVYNKARQASITQELTEIVGASAAV 287

RESULT 12
Q8E8B9
ID Q8E8B9 PRELIMINARY; PRT; 286 AA.
AC Q8E8B9;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DE ATP synthase F1, gamma subunit.
GN Name=atpG; OrderedLocusNames=SO4748;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCB1_TaxID=70863;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsaplin A., Scott J., Beanan M.J.,
RA Brinkac L.W., Dougherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,

```

```

RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Utterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Nealeon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015907; AAN57707.1; -.
DR HSSP; P00837; 1FS0.
DR TIGR; SO4748; -.
DR GO; GO:0015020; C:membrane; IEA.
DR GO; GO:0015469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro; IPR00131; ATPase_gamma.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; PR00126; ATPSEGAMMA.
DR TIGRFAMs; TIGR01146; ATPsyn_Figamma; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; 1.
DR Complete proteome.
DR GO; GO:0015469; C:proton-transporting two-sector ATPase complex; IEA.
DR EMBL; AP005343; BAC96016.1; -.
DR GO; GO:0015020; C:membrane; IEA.
DR GO; GO:0015469; C:proton-transporting two-sector ATPase complex; IEA.

White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
Utterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
Venter J.C., Nealeon K.H., Fraser C.M.;
"Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis.";
Nat. Biotechnol. 20:1118-1123(2002).
EMBL; AE015907; AAN57707.1; -.
HSSP; P00837; 1FS0.
TIGR; SO4748; -.
GO; GO:0015020; C:membrane; IEA.
GO; GO:0015469; C:proton-transporting two-sector ATPase complex; IEA.
GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
GO; GO:0046961; F:hydrogen-transporting ATP synthase activity; IEA.
GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
InterPro; IPR00131; ATPase_gamma.
Pfam; PF00231; ATP-synt; 1.
PRINTS; PR00126; ATPSEGAMMA.
TIGRFAMs; TIGR01146; ATPsyn_Figamma; 1.
PROSITE; PS00153; ATPASE_GAMMA; 1.
Complete proteome.
GO; GO:0015469; C:proton-transporting two-sector ATPase complex; IEA.
EMBL; AP005343; BAC96016.1; -.
GO; GO:0015020; C:membrane; IEA.
GO; GO:0015469; C:proton-transporting two-sector ATPase complex; IEA.

```

```

DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATP synthase activity; IEA.
DR InterPro; IPR00131; ATPase_gamma.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; PR00126; ATPSEGAMMA.
DR TIGRFAMs; TIGR01146; ATPsyn_Figamma; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; 1.
DR SEQUENCE 288 AA; 31710 MW; 1143F7A111A9703C CRC64;

Query Match 68.6%; Score 980.5; DB 2; Length 288;
Best Local Similarity 64.0%; Pred. No. 2.1e-56;
Matches 185; Conservative 58; Mismatches 45; Indels 1; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASVKRKTQERMSSSRPYSETIRNVISHVSKAT 60
Db 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASVKRKTQERMSSSRPYSETIRNVISHVSKAT 60
QY 61 IGYKHPFLVDREVKKVGMIVSTDRGLCGGLNVLNFKTVLNMKEWKEKDVSVOLSLIGS 120
Db 61 LEYRHPFLVDREVKKVGMIVSTDRGLCGGLNVLNFKTVLNMKEWKEKDVSVOLSLIGS 120
QY 121 KSNIPFQSLGKILTDQSGIGTPTSPVEQLIGSVNSMIDAYKKGEVDVYLVYVNFINTMS 180
Db 121 KATAPFKHGAKVAQVSGGLGDSPLSDIGSVNLEKYDEGLDRLYLVFNFVNTMV 180
QY 181 QKPVLEKLIPLPELNDDELGERKQVWDYIYEPDAKVLNDLLVLYLESQVYQAAVENLAS 240
Db 181 QKPTVLEKLIPLPELNDDELGERKQVWDYIYEPDAKVLNDLLVLYLESQVYQAAVENLAS 240
QY 241 EQAARWAMKAATDNAGNLINELQVYNKARQASITTNELNEIVAGAAAI 289
Db 240 EQAARWAMKAATDNAGNLINELQVYNKARQASITTNELNEIVAGAAAI 289

RESULT 14
Q8DDG9 PRELIMINARY; PRT; 288 AA.
AC Q8DDG9;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DE ATP synthase F1, Gamma subunit.
GN OrderedLocusNames=V11020;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Cho Y.H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL EMBL; AE016800; AAC09508.1; -.
DR HSSP; P00837; 1FS0.
DR GO; GO:0015020; C:membrane; IEA.
DR GO; GO:0015469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro; IPR00131; ATP-synt; 1.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; PR00126; ATPSEGAMMA.
DR TIGRFAMs; TIGR01146; ATPsyn_Figamma; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; 1.
DR Complete proteome.
DR SEQUENCE 288 AA; 31710 MW; 1143F7A111A9703C CRC64;

Query Match 68.6%; Score 980.5; DB 2; Length 288;
Best Local Similarity 64.0%; Pred. No. 2.1e-56;
Matches 185; Conservative 58; Mismatches 45; Indels 1; Gaps 1;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 09:48:14 ; Search time 39 Seconds
(without alignments)
491.433 Million cell updates/sec

Title: US-09-545-199F-4

Perfect score: 1429

Sequence: 1 MAGAKEIRTKIASVKSTQKI.....ARQASITNELNEIVAGAAAI 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1429	100.0	289	4	US-09-809-665A-4
2	1132.5	79.3	288	4	US-09-809-665A-133
3	1077.5	75.4	288	4	US-09-809-665A-167
4	1004	70.3	291	4	US-09-543-681A-7449
5	978	68.4	297	4	US-09-489-039A-10616
6	926.5	64.8	291	4	US-09-252-991A-19080
7	894	62.6	292	4	US-09-328-352-4894
8	885	61.9	309	4	US-09-540-236-2848
9	530.5	37.1	304	4	US-09-134-000C-5902
10	524.5	36.7	289	4	US-09-254-504-4
11	510.5	35.7	300	4	US-09-107-532A-7284
12	504.5	35.3	289	4	US-09-254-504-8
13	501.5	35.1	292	4	US-09-583-110-3385
14	493	34.5	295	3	US-09-134-001C-4242
15	464	32.5	272	4	US-09-710-279-1812
16	460.5	32.2	293	4	US-09-254-504-12
17	440	30.8	308	4	US-08-311-731A-148
18	231.5	16.2	281	4	US-09-861-451A-22
19	179	12.5	183	4	US-09-248-796A-17416
20	143.5	10.0	139	4	US-09-513-995C-4941
21	136.5	9.6	136	4	US-09-270-767-32542
22	136.5	9.6	136	4	US-09-270-767-47759
23	108.5	7.6	383	4	US-09-328-352-7081
24	102.5	7.2	10182	3	US-09-134-001C-3159
25	102	7.1	386	2	US-08-455-968E-7
26	100	7.0	721	4	US-09-248-796A-15107
27	97	6.8	289	4	US-09-107-532A-4402

28	95	6.6	430	4	US-09-328-352-5624	Sequence 5624, Ap
29	94	6.6	487	4	US-09-107-532A-5731	Sequence 5731, Ap
30	93.5	6.5	718	3	US-08-973-005A-12	Sequence 12, Appl
31	92	6.4	315	3	US-08-793-035-9	Sequence 9, Appl
32	92	6.4	315	3	US-08-793-035-10	Sequence 10, Appl
33	92	6.4	775	1	US-07-603-133B-16	Sequence 16, Appl
34	92	6.4	786	4	US-09-248-796A-15288	Sequence 15288, A
35	91.5	6.4	635	4	US-09-538-092-1	Sequence 1, Appl
36	91.5	6.4	2349	4	US-09-538-092-914	Sequence 914, Ap
37	91.5	6.4	2710	1	US-08-480-604A-6	Sequence 6, Appl
38	91.5	6.4	2710	2	US-08-405-496A-6	Sequence 6, Appl
39	91.5	6.4	2710	3	US-08-915-136-6	Sequence 6, Appl
40	91.5	6.4	2710	3	US-08-957-310-6	Sequence 6, Appl
41	91.5	6.4	2710	4	US-10-011-366-6	Sequence 6, Appl
42	91.5	6.4	2710	4	US-09-084-517-6	Sequence 6, Appl
43	91.5	6.4	3878	4	US-09-914-259-11	Sequence 11, Appl
44	91	6.4	444	4	US-09-248-796A-19173	Sequence 19173, A
45	91	6.4	1010	3	US-09-134-001C-5178	Sequence 5178, Ap

ALIGNMENTS

RESULT 1

US-09-809-665A-4
; Sequence 4, Application US/09809665A
; Patent No. 6790950

; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.

; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 2834/00435

; CURRENT APPLICATION NUMBER: US/09/809,665A
; PRIOR FILING DATE: 2001-03-15

; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10

; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09

; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06

; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Pasteurella multocida

US-09-809-665A-4

Query Match 100.0%; Score 1429; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 2.4e-136;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGAKEIRTKIASVKSTQKI	KAMEMVAASVKRKTQREMSRRPYSETIRNVISHVSKAT	60
DB	1	MAGAKEIRTKIASVKSTQKI	KAMEMVAASVKRKTQREMSRRPYSETIRNVISHVSKAT	60
QY	61	IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNNLFTVLNEMKEKDVSVQLSLIGS	120	
DB	61	IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNNLFTVLNEMKEKDVSVQLSLIGS	120	
QY	121	KSINFFQSLGKILITQDSGIGDTPSVQQLIGSVNSMIDAYKKEGVDVVYLVYKFNITMS	180	
DB	121	KSINFFQSLGKILITQDSGIGDTPSVQQLIGSVNSMIDAYKKEGVDVVYLVYKFNITMS	180	
QY	181	OKPVLEKLIPLPELNDDELGERKQVNDYIYEPDAKVLNDLLVRYLESQVYQAAVENLAS	240	
DB	181	OKPVLEKLIPLPELNDDELGERKQVNDYIYEPDAKVLNDLLVRYLESQVYQAAVENLAS	240	
QY	241	EQAAARMVAMKAATDNAGNLINELQLVYKARQASITNELNEIVAGAAAI	289	
DB	241	EQAAARMVAMKAATDNAGNLINELQLVYKARQASITNELNEIVAGAAAI	289	

RESULT 2

US-09-809-665A-133
 ; Sequence 133, Application US/09809665A
 ; Patent No. 6790950
 ; GENERAL INFORMATION:
 ; APPLICANT: Lowery E., David, et al.
 ; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
 ; FILE REFERENCE: 28341/00435
 ; CURRENT APPLICATION NUMBER: US/09/809,665A
 ; PRIOR FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 60/153,453
 ; PRIOR FILING DATE: 1999-09-10
 ; PRIOR APPLICATION NUMBER: 60/128,689
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 09/545,199
 ; PRIOR FILING DATE: 2000-04-06
 ; NUMBER OF SEQ ID NOS: 197
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 133
 ; LENGTH: 288
 ; TYPE: PRT
 ; ORGANISM: Actinobacillus pleuropneumoniae
 US-09-809-665A-133

Query Match 79.3%; Score 1132.5; DB 4; Length 288;
 Best Local Similarity 74.4%; Pred. No. 2.4e-106;
 Matches 215; Conservative 45; Mismatches 28; Indels 1; Gaps 1;

QY 1 MAGAKEIRTKIASVSTQKITKAMENVAASRKRTQERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVSTQKITKAMENVAASRKRTQERMSSSRPYSETIRNVISHVSKAT 60
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKTVLNMKEWKEKQVSVLSLIGS 120
 DB 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKTVLNMKEWKEKQVSVLSLIGS 120
 QY 121 KSNFPQSLGKILTDQSGIGTPEVEOLIGSVNSMIDAYKKGVDVYVLYVYNKFNITMS 180
 DB 121 KGVSPYQNLGNVRSQVTLGDNPEMERIVGAVNEMINAFNGEVDVAVYVYNKFNITMS 180
 QY 181 QKPVLKLIPLPELNDDELGERKQVWDYIPEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QKPVLKLIPLPELNDDELGERKQVWDYIPEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 QY 241 EQAARMVAMKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 289
 DB 240 EQAARMVAMKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 288

RESULT 3

US-09-809-665A-167
 ; Sequence 167, Application US/09809665A
 ; Patent No. 6790950
 ; GENERAL INFORMATION:
 ; APPLICANT: Lowery E., David, et al.
 ; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
 ; FILE REFERENCE: 28341/00435
 ; CURRENT APPLICATION NUMBER: US/09/809,665A
 ; PRIOR FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 60/153,453
 ; PRIOR FILING DATE: 1999-09-10
 ; PRIOR APPLICATION NUMBER: 60/128,689
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 09/545,199
 ; PRIOR FILING DATE: 2000-04-06
 ; NUMBER OF SEQ ID NOS: 197
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 167
 ; LENGTH: 288
 ; TYPE: PRT
 ; ORGANISM: Pasteurella (Mannheimia) haemolytica
 US-09-809-665A-167

Query Match 75.4%; Score 1077.5; DB 4; Length 288;
 ; Sequence 167, Application US/09809665A
 ; Patent No. 6790950

Best Local Similarity 72.0%; Pred. No. 8.9e-101;
 Matches 208; Conservative 46; Mismatches 34; Indels 1; Gaps 1;
 QY 1 MAGAKEIRTKIASVSTQKITKAMENVAASRKRTQERMSSSRPYSETIRNVISHVSKAT 60
 DB 1 MAGAKEIRTKIASVSTQKITKAMENVAASRKRTQERMSSSRPYSETIRNVISHVSKAT 60
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKTVLNMKEWKEKQVSVLSLIGS 120
 DB 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKTVLNMKEWKEKQVSVLSLIGS 120
 QY 121 KSNFPQSLGKILTDQSGIGTPEVEOLIGSVNSMIDAYKKGVDVYVLYVYNKFNITMS 180
 DB 121 KGVSPYQNLGNVRSQVTLGDNPEMERIVGAVNEMINAFNGEVDVAVYVYNKFNITMS 180
 QY 181 QKPVLKLIPLPELNDDELGERKQVWDYIPEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 181 QKPVLKLIPLPELNDDELGERKQVWDYIPEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 QY 241 EQAARMVAMKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 289
 DB 240 EQAARMVAMKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 288

RESULT 4

US-09-543-681A-7449
 ; Sequence 7449, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 7449
 ; LENGTH: 291
 ; TYPE: PRT
 ; ORGANISM: Proteus mirabilis
 US-09-543-681A-7449

Query Match 70.3%; Score 1004; DB 4; Length 291;
 Best Local Similarity 67.1%; Pred. No. 2.5e-93;
 Matches 194; Conservative 49; Mismatches 44; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVSTQKITKAMENVAASRKRTQERMSSSRPYSETIRNVISHVSKAT 60
 DB 5 MAGAKEIRTKIASVSTQKITKAMENVAASRKRTQERMSSSRPYSETIRNVISHVSKAT 64
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKTVLNMKEWKEKQVSVLSLIGS 120
 DB 65 LEYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKTVLNMKEWKEKQVSVLSLIGS 124
 QY 121 KSNFPQSLGKILTDQSGIGTPEVEOLIGSVNSMIDAYKKGVDVYVLYVYNKFNITMS 180
 DB 125 KAVSPFASVGVNNGVVTGMDPQLSDLGIVNMLQAYDEGLDKLYVYNKFNITMS 184
 QY 181 QKPVLKLIPLPELNDDELGERKQVWDYIPEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 DB 185 QEPKLVLPPLPGDDDELKES--WDYLYPEPDKTLTLLRYLESQVYQAAVENLAS 242
 QY 241 EQAARMVAMKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAI 289
 DB 243 EQAARMVAMKAATDNAGLNELQVYNKARQASITNELNEIVAGAAAV 291

RESULT 5

US-09-489-039A-10616
 ; Sequence 10616, Application US/09489039A
 ; Patent No. 6610836

; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 10616
 ; LENGTH: 297
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-10616

Query Match 68.4%; Score 978; DB 4; Length 297;
 Best Local Similarity 65.1%; Pred. No. 1.1e-90;
 Matches 188; Conservative 50; Mismatches 49; Indels 2; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKRTQRMSSSRPYSETIRNVISHVSKAT 60
 Db 11 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKRTQRMSSSRPYSETIRNVISHVSKAT 70
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKVLNEMKEKDVSVQSLIGS 120
 Db 71 LEYKHPLEERDVKKVGMIVVSTDRGLCGGLNVLNFKVLNEMKEKDVSVQSLIGS 130
 QY 121 KSNFFQSLGKILTPDSGIGTSPVEQLIGSVNSMIDAYKKGEVDVYVLYVYKKEINTMS 180
 Db 131 KGVSFNSVGGNVVAVQTGMDNPSLSLIGPVKWLQAYDEGRDLKLYVSNKEINTMS 190
 QY 181 QKPVLEKLIPLPDLNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 Db 191 QVTITQLPLPASEADL--KKKSWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 248
 QY 241 EQAARMVAMKAATDNAGNLINELQVYKARQASITNELNEIVAGAAAI 289
 Db 249 EQAARMVAMKAATDNAGNLINELQVYKARQASITNELNEIVAGAAAI 297

RESULT 6
 US-09-252-991A-19080
 ; Sequence 19080, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: 107196.136
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 19080
 ; LENGTH: 291
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-19080

Query Match 64.8%; Score 926.5; DB 4; Length 291;
 Best Local Similarity 62.3%; Pred. No. 1.7e-85;
 Matches 180; Conservative 50; Mismatches 56; Indels 3; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKRTQRMSSSRPYSETIRNVISHVSKAT 60
 Db 6 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKRTQRMSSSRPYSETIRNVISHVSKAT 65
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKVLNEMKEKDVSVQSLIGS 120

Db 66 PEYRHPFVREVKRVGYIVVSSDRGLCGGLNVLNFKVLNEMKEKDVSVQSLIGS 125
 QY 121 KSNFFQSLGKILTPDSGIGTSPVEQLIGSVNSMIDAYKKGEVDVYVLYVYKKEINTMS 180
 Db 126 KQASFFRSFGGNVAAISHLGEBSINDLIGSVKWLMDAYLEGRIQLRFLVSVNKFVNTMT 185
 QY 181 QKPVLEKLIPLPDLNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 240
 Db 186 QKPTVEQLIPLVADDQEL--KHWWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 242
 QY 241 EQAARMVAMKAATDNAGNLINELQVYKARQASITNELNEIVAGAAAI 289
 Db 243 EQAARMVAMKAATDNAGNLINELQVYKARQASITNELNEIVAGAAAI 291

RESULT 7
 US-09-328-352-4894
 ; Sequence 4894, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; PRIOR FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 4894
 ; LENGTH: 292
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-4894

Query Match 62.6%; Score 894; DB 4; Length 292;
 Best Local Similarity 59.7%; Pred. No. 3.4e-82;
 Matches 173; Conservative 56; Mismatches 59; Indels 2; Gaps 2;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASQKRTQRMSSSRPYSETIRNVISHVSKAT 60
 Db 4 MANLAEIRAKVASIKSTQKITRANQVAAASQKRAQERMAQGRFYADNMRVIAHLVQAN 63
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVLNFKVLNEMKEKDVSVQSLIGS 120
 Db 64 PEYKRYVMDRVPKRVGYIVVSSDRGLAGGLNVLNFKVYQHVKAQQEQSIEVQFALIGQ 123
 QY 121 KSNFFQSLGKILTPDSGIGTSPVEQLIGSVNSMIDAYKKGEVDVYVLYVYKKEINTMS 180
 Db 124 KAVSEFKYGGKVLGATQTGDAFSLQTLGTSVQVMDAFDKGELDRIVLVSNQFVNAWT 183
 QY 181 QKPVLEKLIPLPDLNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 239
 Db 184 QKPVLEKLIPLPDLNDELGERKQVWDYIYEPDAKVLNLLVRYLESQVYQAAVENLAS 242
 QY 240 EQAARMVAMKAATDNAGNLINELQVYKARQASITNELNEIVAGAAAI 289
 Db 243 EQAARMVAMKAATDNAGNLINELQVYKARQASITNELNEIVAGAAAI 292

RESULT 8
 US-09-540-236-2848
 ; Sequence 2848, Application US/09540236
 ; Patent No. 6673910
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
 ; FILE REFERENCE: 2709.2005-001
 ; CURRENT APPLICATION NUMBER: US/09/540,236
 ; PRIOR FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 3840
 ; SEQ ID NO 2848
 ; LENGTH: 309
 ; TYPE: PRT


```

1  ZIP: 02354
2
3  COMPUTER READABLE FORM:
4
5  MEDIUM TYPE: CD/ROM ISO9660
6
7  COMPUTER: PC
8
9  OPERATING SYSTEM: <Unknown>
10
11 SOFTWARE: ASCII
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/09/107,532A
16
17 FILING DATE: 30-Jun-1998
18
19 PRIOR APPLICATION DATA:
20
21 APPLICATION NUMBER: 60/085,598
22
23 FILING DATE: 14 May 1998
24
25 APPLICATION NUMBER: 60/051571
26
27 FILING DATE: July 2, 1997
28
29 ATTORNEY/AGENT INFORMATION:
30
31 NAME: Ariniello, Pamela Deneké
32
33 REGISTRATION NUMBER: 40,489
34
35 REFERENCE/DOCKET NUMBER: GTC-012
36
37 TELECOMMUNICATION INFORMATION:
38
39 TELEPHONE: (781)893-5007
40
41 TELEFAX: (781)893-8277
42
43 INFORMATION FOR SEQ ID NO: 7284:
44
45 SEQUENCE CHARACTERISTICS:
46
47 LENGTH: 300 amino acids
48
49 TYPE: amino acid
50
51 TOPOLOGY: linear
52
53 MOLECULE TYPE: protein
54
55 HYPOTHETICAL: YES
56
57 ORIGINAL SOURCE:
58
59 ORGANISM: Enterococcus faecium
60
61 FEATURE:
62
63 NAME/KEY: misc_feature
64
65 LOCATION: (B) LOCATION 1...300
66
67 SEQUENCE DESCRIPTION: SEQ ID NO: 7284:
68
69 US-09-107-532A-7284

```

Query Match	35.7%	Score	510.5	DB	4	Length	300
Best Local Similarity	36.3%	Pred. No.	2.4e-43				
Matches	109	Conservative	69	Mismatches	107	Indels	15
						Gaps	3
QY	2	AGAEIRTKASVSTOKITKAMEMVAASOKRKTQRMSSSRPYSETIRNVISHVS	---	57			
Db	3	ASLNEIKTRIASTKTKTSQITIRAMQMVSAASKLTKSEASSQKFQYATKVREIVTHLTATQL	62				
QY	58	-----KATIGYKHFFLDVREVKVGMIVUSTDRGLCGGLNVNLFKTVLNEMKEWKEK	109				
Db	63	NDIASENPRGDINT-NSMLLSRPVKTKGYIVITADGGLVGGYNSILKOTYSILIEEDHQ	121				
QY	110	DVSQVLSLIGSKSINFFQSGLIKILTQDSGIGDTPSPVEQLIGSVNSMIDAYKKGVEVDVWY	169				
Db	122	PDDYVMTAIGTGADFFKAGIINLAYELRNLSDPSPFDEVRKIVNMATWYQNEVFEDELY	181				
QY	170	LVYANKFTNWSQRPVLEKLPFLDELONDELGERKQWYDIYEPDAKVLLDNLVRYLESQ	239				
Db	182	VCYNHHINSUTSQRPVEKMLPISDLPQSEATTFEQ--EYIPEPSKEETLSQLLQFYAESL	239				
QY	230	VYQAAVENLASQAAARVMKAAATDNAGLNLINELQLVNFKARQASITINELNEIVAGAAAI	289				
Db	240	IYGAIVDAKTAEEHAAGTAKMTATDNAANIIDELTVSNRAQOAGTIOETIEIVAGASAL	299				

```

RESULT 12
US-09-254-504-8
; Sequence 8, Application US/09254504
; Patent No. 6511836
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A method of improving the production of biomass
; TITLE OF INVENTION: or a desired product from a cell
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

/ / OPERATING SYSTEM: PC-DOS/MS-DOS
/ / SOFTWARE: Patent In Release #1.0,
/ / CURRENT APPLICATION DATA:
/ / APPLICATION NUMBER: US/09/254,504
/ / FILING DATE:
/ / PRIOR APPLICATION DATA:
/ / APPLICATION NUMBER: DK 963/96
/ / FILING DATE: 06-SEP-1996
/ / INFORMATION FOR SEQ ID NO: 8:
/ / SEQUENCE CHARACTERISTICS:
/ / LENGTH: 289 amino acids
/ / TYPE: amino acid
/ / TOPOLOGY: linear
/ / MOLECULE TYPE: protein
/ / US-09-254-504-8

```

Query Match	35.3%;	Score	504.5;	DB	4;	Length	289;
Best Local Similarity	36.6%;	Pred. No.	9e-43;				
Matches	106;	Conservative	70;	Mismatches	109;	Indels	5; Gaps 3;
QY	2	AGAKEIRTKIASVKSTOKITKAMEMVAASQKRTQERMSSRPYSYETIRNVISH-VSKAT	60				
Db	3	ASLNEIKTKIASTKTKTSQITGAQMWASAALKQAESHAKAFOIYAERVKRITDVLVSSDK	62				
QY	51	IGYKHPFLVDREYKKYGMIVVSTDRGLCGGLNVNLEKTVLNEWKKEWZDVLSQLTGS	120				
Db	63	EPANPMIMIGREVKTKTYLITSDRLGVGYNSILKSVNMTIRKRPANESYTTILAGG	122				
QY	121	KSINFOSLGIKITDQSGIGDTPSPVEQLIGSVNSMIDAYKKGEVDVYVLYVYKKEINTWS	180				
Db	123	TGADFFGASNVKSFVLGCFSPDQNFEEVRAIVTEAVTEYQAEFPDELVCVCYNHVNLSLV	182				
QY	181	QKPVLEKLIPLPELDNDELGERQVND-IYYPDPAKVLLDNLIVLYBSQVYQAAVENLIA	239				
Db	183	SEASMEKXLP---IFFEASCCQPPFFTELEPDCETILNQLLPPYASMSIYGSIVDAKT	239				
QY	240	SEQAARVAMKAATDNAGNLINELQIVYNKARQASITNELNEIVAGAAAI	289				
Db	240	AEHAAGTAMRTATDNAAHSYINDLTTCYNRAQASITQITETIVAGAGAL	289				

```

RESULT 13
US-09-583-110-3385
; Sequence 3385, Application US/09583110
; Patent NO. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lyvin Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Anti-
; TITLE OF INVENTION: Pneumoniae for Diagn
; FILE REFERENCES: PATHO0-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3385
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3385

```

```

Query Match      35.1%; Score 501.5; DB 4; Length 292;
Best Local Similarity 35.3%; Pred.No.1.8e-42;
Matches 101; Conservative 78; Mismatches 104; Indels 3; Gaps 2;

6 EIRKTASVASTQKITKAEWMAASKORITQRMSSRPYSETIRNVISHV--SKATIGY 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 DIKTKASTNTSQIINAVQWVSAAKLGEAEAAFNFOVYAKVKRLITDILHNGAGAS 66
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 09:50:44 ; Search time 140 Seconds
(without alignments)
723.159 Million cell updates/sec

Title: US-09-545-199F-4

Perfect score: 1429

Sequence: 1 MAGAKEIRTKIASVKSTQKI.....ARQASITNELNEIVAGAAAI 289

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1429	100.0	289	11	US-09-809-665A-4
2	1132.5	79.3	288	11	US-09-809-665A-133
3	1077.5	75.4	288	11	US-09-809-665A-167
4	1007	70.5	287	14	US-10-369-493-21244
5	1002	70.1	287	14	US-10-369-493-436
6	991	69.3	287	9	US-09-741-669-337
7	991	69.3	287	14	US-10-369-493-23613
8	904.5	63.3	286	14	US-10-369-493-13845
9	891	62.4	289	14	US-10-369-493-8843
10	885.5	62.0	294	14	US-10-369-493-19599
11	838	58.6	287	14	US-10-369-493-15422
12	838	58.6	287	14	US-10-369-493-15790
13	838	58.6	287	14	US-10-369-493-16172

14	812.5	56.9	292	14	US-10-369-493-4487	Sequence 4487, Ap
15	812.5	56.9	292	14	US-10-369-493-7246	Sequence 7246, Ap
16	809	56.6	287	14	US-10-369-493-9371	Sequence 9371, Ap
17	795	55.6	287	14	US-10-369-493-17565	Sequence 17565, A
18	787	55.1	287	14	US-10-369-493-9279	Sequence 9279, Ap
19	585.5	41.0	281	14	US-10-369-493-9902	Sequence 9902, Ap
20	567.5	39.7	293	14	US-10-369-493-8913	Sequence 8913, Ap
21	536.5	37.5	285	14	US-10-369-493-17454	Sequence 17454, A
22	534.5	37.4	315	14	US-10-369-493-20262	Sequence 20262, A
23	533.5	37.3	287	14	US-10-369-493-23329	Sequence 23329, A
24	529	37.0	286	14	US-10-369-493-16552	Sequence 16552, A
25	524.5	36.7	289	14	US-10-310-630-4	Sequence 4, Appli
26	523.5	36.6	289	14	US-10-369-493-18487	Sequence 18487, A
27	509.5	35.7	315	14	US-10-369-493-18941	Sequence 8, Appli
28	504.5	35.3	289	14	US-10-310-630-8	Sequence 18941, A
29	504.5	35.3	314	14	US-10-369-493-2612	Sequence 2612, Ap
30	502	35.1	294	14	US-10-369-493-19549	Sequence 19549, A
31	501.5	35.1	292	16	US-10-474-776-690	Sequence 690, App
32	481	33.7	315	14	US-10-369-493-20871	Sequence 20871, A
33	469.5	32.9	290	14	US-10-369-493-10691	Sequence 10691, A
34	469	32.6	360	15	US-10-767-701-47119	Sequence 47119, A
35	466.5	32.6	301	15	US-10-335-977-5746	Sequence 5746, Ap
36	463	32.4	357	15	US-10-425-114-67588	Sequence 67588, A
37	463	32.4	359	17	US-10-425-115-360784	Sequence 360784,
38	461.5	32.3	325	9	US-09-738-626-4841	Sequence 4841, Ap
39	461.5	32.3	325	16	US-10-781-014-776	Sequence 776, App
40	460.5	32.2	290	14	US-10-369-493-10114	Sequence 10114, A
41	460.5	32.2	293	14	US-10-310-630-12	Sequence 12, Appl
42	459.5	32.2	292	14	US-10-369-493-11433	Sequence 11433, A
43	459.5	32.2	292	14	US-10-369-493-14233	Sequence 14233, A
44	459.5	32.2	292	14	US-10-369-493-14452	Sequence 14452, A
45	459.5	32.2	292	14	US-10-369-493-14863	Sequence 14863, A

ALIGNMENTS

RESULT 1
US-09-809-665A-4
; Sequence 4, Application US/09809665A
; Publication No. US20040110268A1
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28341/00435
; CURRENT APPLICATION NUMBER: US/09/809,665A
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-809-665A-4

Query Match	100.0%	Score 1429;	DB 11;	Length 289;
Best Local Similarity	100.0%	Pred. No. 2.1e-118;		
Matches 289;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAGAKEIRTKIASVKSTQKI	TKAMENVAASKMKTKQERMSRRPYSETIRNVISHVSKAT	60
DB	1	MAGAKEIRTKIASVKSTQKI	TKAMENVAASKMKTKQERMSRRPYSETIRNVISHVSKAT	60
QY	61	IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTLVNLNEMKEWKEKDVSVQLSLIGS	120	
DB	61	IGYKHPFLVDREVKVGMIVVSTDRGLCGGLNVNFKTLVNLNEMKEWKEKDVSVQLSLIGS	120	

QY 121 KSNFFQSLGKILTDQSGIGDTPSVQLIGSVNSMIDAYKKGEVDVYLVNKFINTMS 180
 |||||
 Db 121 KSNFFQSLGKILTDQSGIGDTPSVQLIGSVNSMIDAYKKGEVDVYLVNKFINTMS 180
 |||||
 QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLDDNLLVRYLESQVYQAAVENLAS 240
 |||||
 Db 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLDDNLLVRYLESQVYQAAVENLAS 240
 |||||
 QY 241 EQARWVAMKAATDNAGNLINELQLVNKAQASITNELNEIVAGAAAI 289
 |||||
 Db 241 EQARWVAMKAATDNAGNLINELQLVNKAQASITNELNEIVAGAAAI 289
 |||||

RESULT 2

US-09-809-665A-133
 ; Sequence 133, Application US/09809665A
 ; Publication No. US20040110268A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lowery E., David, et al.
 ; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
 ; FILE REFERENCE: 28341/00435
 ; CURRENT APPLICATION NUMBER: US/09/809,665A
 ; CURRENT FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 60/153,453
 ; PRIOR FILING DATE: 1999-09-10
 ; PRIOR APPLICATION NUMBER: 60/128,689
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 09/545,199
 ; PRIOR FILING DATE: 2000-04-06
 ; NUMBER OF SEQ ID NOS: 197
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 133
 ; LENGTH: 288
 ; TYPE: PRT
 ; ORGANISM: Actinobacillus pleuropneumoniae
 US-09-809-665A-133

Query Match 79.3%; Score 1132.5; DB 11; Length 288;
 Best Local Similarity 74.4%; Pred. No. 4.4e-92;
 Matches 215; Conservative 45; Mismatches 28; Indels 1; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKMRKTKQERMSSSRPYSETIRNVISHVSKAT 60
 |||||
 Db 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKMRKTKQERMSSSRPYSETIRNVISHVSKAT 60
 |||||
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 120
 |||||
 Db 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 120
 |||||
 QY 121 KSNFFQSLGKILTDQSGIGDTPSVQLIGSVNSMIDAYKKGEVDVYLVNKFINTMS 180
 |||||
 Db 121 KSVFYQNLGNVRSQVTLGDNPEMERIVGAVNEVINAFRNGEVDVAVYNRFINTMS 180
 |||||
 QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLDDNLLVRYLESQVYQAAVENLAS 240
 |||||
 Db 181 QKPVIAQLPLPKLDDDEL-DTKGSMWDYIYEPNPQVLLDRLVRYLETQVYQAAVDNLAS 239
 |||||
 QY 241 EQARWVAMKAATDNAGNLINELQLVNKAQASITNELNEIVAGAAAI 289
 |||||
 Db 240 EQARWVAMKAATDNAGTLIDELQLVNKAQASITNELNEIVAGAAAI 288
 |||||

RESULT 3

US-09-809-665A-167
 ; Sequence 167, Application US/09809665A
 ; Publication No. US20040110268A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lowery E., David, et al.
 ; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
 ; FILE REFERENCE: 28341/00435
 ; CURRENT APPLICATION NUMBER: US/09/809,665A
 ; CURRENT FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 60/153,453

; PRIOR FILING DATE: 1999-09-10
 ; PRIOR APPLICATION NUMBER: 60/128,689
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 09/545,199
 ; PRIOR FILING DATE: 2000-04-06
 ; NUMBER OF SEQ ID NOS: 197
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 167
 ; LENGTH: 288
 ; TYPE: PRT
 ; ORGANISM: Pasteurella (Mannheimia) haemolytica
 US-09-809-665A-167

Query Match 75.4%; Score 1077.5; DB 11; Length 288;
 Best Local Similarity 72.0%; Pred. No. 3.4e-87;
 Matches 208; Conservative 46; Mismatches 34; Indels 1; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKMRKTKQERMSSSRPYSETIRNVISHVSKAT 60
 |||||
 Db 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKMRKTKQERMSSSRPYSETIRNVISHVSKAT 60
 |||||
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 120
 |||||
 Db 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 120
 |||||
 QY 121 KSNFFQSLGKILTDQSGIGDTPSVQLIGSVNSMIDAYKKGEVDVYLVNKFINTMS 180
 |||||
 Db 121 KSVFYQNLGNVRSQVTLGDNPEMERIVGAVNEVINAFRNGEVDVAVYNRFINTMS 180
 |||||
 QY 181 QKPVLEKLIPLPELDNDELGERKQVWDYIYEPDAKVLDDNLLVRYLESQVYQAAVENLAS 240
 |||||
 Db 181 QKPTVQQLPLPALENDL-EQTGSMWDYIYEPNPQALLDRLVRYLESQVYQAAVDNLAS 239
 |||||
 QY 241 EQARWVAMKAATDNAGNLINELQLVNKAQASITNELNEIVAGAAAI 289
 |||||
 Db 240 EQARWVAMKAATDNAGNLINELQLVNKAQASITNELNEIVAGAAAI 288
 |||||

RESULT 4

US-10-369-493-21244
 ; Sequence 21244, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 21244
 ; LENGTH: 287
 ; TYPE: PRT
 ; ORGANISM: Xenorhabdus nematophilus
 US-10-369-493-21244

Query Match 70.5%; Score 1007; DB 14; Length 287;
 Best Local Similarity 67.5%; Pred. No. 6e-81;
 Matches 195; Conservative 47; Mismatches 45; Indels 2; Gaps 1;
 QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKMRKTKQERMSSSRPYSETIRNVISHVSKAT 60
 |||||
 Db 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKMRKTKQERMSSSRPYSETIRNVISHVSKAT 60
 |||||
 QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 120
 |||||
 Db 61 LETRHPFLDREVKKVGMIVVSTDRGLCGGLNNLFTVLNEMKEWKEKDVSVQLSLIGS 120
 |||||


```

QY 121 KGINFOSLGKILTQDSGIGTPSVEQLIGSVNSMIDAYKKGVDDVVVLYVYNKFTNTMS 180
Db 121 KGVSPFNSVGGNVAAVQVTGMDNPSLSLIGPVKVMQAVDEGRDLKLYIVSNKFTNTMS 180
QY 181 QKPVLEKLIPLPELDNDDELGERKQVWDYIYEPDAKVLNDLLVRYLESQVYQAAVENLAS 240
Db 181 QVPTISQLPLPASDDDL--KHSWDLVYEPDPKALLDITLLRRYVESQVYQGVWENLAS 238
QY 241 EQAARWAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
Db 239 EQAARWAMKAATDNGSLIEXLQVYNKARQASITQELTEIVGAAAV 287

RESULT 8
US-10-369-493-13845
; Sequence 13845, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13845
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13845

```

```

Query Match 63.3%; Score 904.5; DB 14; Length 286;
Best Local Similarity 61.2%; Pred. No. 7.5e-72;
Matches 177; Conservative 49; Mismatches 60; Indels 3; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKRKTKQRMSSSRPYSETIRNVISHVSKAT 60
Db 1 MAGAKEIRSKIASIKSTQKITKAMEKVAASKRKTKQRMSSSRPYSETIRNVISHVSKAT 60
QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLPKTVLNEMKWKEDVSVQLSLIGS 120
Db 61 PEYRHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLPKTVLNEMKWKEDVSVQLSLIGS 120
QY 121 KGINFOSLGKILTQDSGIGTPSVEQLIGSVNSMIDAYKKGVDDVVVLYVYNKFTNTMS 180
Db 121 KGAAFFRNFGNVAAISHLGEPSINDLIGSVKVMQAVDEGRDLKLYIVSNKFTNTMT 180
QY 181 QKPVLEKLIPLPELDNDDELGERKQVWDYIYEPDAKVLNDLLVRYLESQVYQAAVENLAS 240
Db 181 QCPVEQILPLVATPDQEL--KHHWDLVYEPDAKELLDMVRYVESQVYQAAVENNA 237
QY 241 EQAARWAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAAAI 289
Db 238 EQAARNIAMKNATDNAGDLISDLQLVYNKARQAAITQETSEIVGAAAV 286

RESULT 9
US-10-369-493-8843
; Sequence 8843, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

```

```

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8843
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Ralstonia metallidurans
US-10-369-493-8843

```

```

Query Match 62.4%; Score 891; DB 14; Length 289;
Best Local Similarity 59.2%; Pred. No. 1.2e-70;
Matches 171; Conservative 57; Mismatches 59; Indels 2; Gaps 1;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKRKTKQRMSSSRPYSETIRNVISHVSKAT 60
Db 1 MAGTKEIRTKIKSVQNTKTKAMEMVAASKRKTKQRMSSSRPYSETIRNVISHVSKAT 60
QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLPKTVLNEMKWKEDVSVQLSLIGS 120
Db 61 PEYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLPKTVLNEMKWKEDVSVQLSLIGS 120
QY 121 KGINFOSLGKILTQDSGIGTPSVEQLIGSVNSMIDAYKKGVDDVVVLYVYNKFTNTMS 180
Db 121 KGMQFLGRIGAKVSVHVQVLDGTPHLEKLGIAIKVQLDAYTNGEVDVAVLYATKFTNTMK 180
QY 181 QKPVLEKLIPL--PELDNDDELGERKQVWDYIYEPDAKVLNDLLVRYLESQVYQAAVENL 238
Db 181 QEPWVEQLPLAADKLQSQTEDEKRAYSDYIYEPDAQTVVEELLVRYVEALVYQAAENM 240
QY 239 ASEQAARWAMKAATDNAGNLINELQVYNKARQASITNELNEIVAGAA 287
Db 241 ASEQAARWAMKAASDNKNVIGELQVYNKTRQAAITKELSELVSGAA 289

```

```

RESULT 10
US-10-369-493-19599
; Sequence 19599, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19599
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Nitrosomonas europaea
US-10-369-493-19599

```

```

Query Match 62.0%; Score 885.5; DB 14; Length 294;
Best Local Similarity 56.1%; Pred. No. 3.8e-70;
Matches 165; Conservative 67; Mismatches 57; Indels 5; Gaps 2;

QY 1 MAGAKEIRTKIASVKSTQKITKAMEMVAASKRKTKQRMSSSRPYSETIRNVISHVSKAT 60
Db 1 MPSSREIRNKIKSVQNTKTKAMEMVAASKRKTKQRMSSSRPYSETIRNVISHVSKAT 60
QY 61 IGYKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLPKTVLNEMKWKEDVSVQLSLIG 119

```

Db 61 VEYRHPFLSRDSVKRGVIVVTSKGLCGGLNTNVLRRALNEIRTWETEGNHVDAACIG 120
QY 120 SKSINFFQSLGKILTDQSGIGDTPSVEQLIGSVNSMIDAYKKGEDVVDVYLVNKFINTM 179
Db 121 NKGLGFWSLGTQVTSQVTLGSDAPNMEELIGAVKVIDATGEGOLDRVYIFYNRFINTM 180
QY 180 SKQPVLEKLIPIPE-----LDNDELGERKQVMDYIYEPDAKVLNLLVRYLESQVQAAY 235
Db 181 KQPVNVEQLLPTDDRISSEDGEARPTRAPMDYIYEPDAKVLNLLVRYLESQVQAAY 240
QY 236 ENLASQARPMVAMKAATDNAGLNLQLVNKAQASITNELNEIVAGAAI 289
Db 241 ENWASEQSRPMVAMKAASDNAGNLDELTLIYNKSRQAATQEISEIVSGAAV 294

RESULT 11
US-10-369-493-15422
; Sequence 15422, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15422
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15422

Query Match 58.6%; Score 838; DB 14; Length 287;
Best Local Similarity 53.9%; Pred. No. 6.1e-66;
Matches 158; Conservative 69; Mismatches 56; Indels 10; Gaps 3;
QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASQKRTQERMSSSRPYSETIRNVISHVSKAT 60
Db 1 MAGGREIKTKISVQNTKRVTRALEMVSASKIRKAQERMKTSRPAQAMQVIGHLAQAS 60
QY 61 IGYKHPFLVDRE-VKKVGMIVVSTDRLCGGLNVNLFKTVLNMKEWKEKDVSVQLSLIG 119
Db 61 TDYQHPFLVEREQVRGVYIVISSDRGLAGLNNLFRKMLGEVRPMQDKGAEIDVVVTIG 120
QY 120 SKSINFFQSLGKILTDQSGIGDTPSVEQLIGSVNSMIDAYKKGEDVVDVYLVNKFINTM 179
Db 121 OKASAFFRIKVMVGSVTHLGDSPHIEQLVGVIKWMLDAFTGKVDVRLVYVNRFNVTM 180
QY 180 SKQPVLEKLIPIPELDNDELGERKQV---WDYIYEPDAKVLNLLVRYLESQVQAAY 236
Db 181 TQKASFQELLP-----AAEHKVAHHDWDYIYEPDAATVLEHVMTRYTESLVYQAVLE 234
QY 237 NLASEQARPMVAMKAATDNAGLNLQLVNKAQASITNELNEIVAGAAI 289
Db 235 NVASEHAARPMVAMKAASDNANKMIGTLQLVNKAQQAATQEISEIVSGAAV 287

RESULT 12
US-10-369-493-15790
; Sequence 15790, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15790
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15790

Query Match 58.6%; Score 838; DB 14; Length 287;
Best Local Similarity 53.9%; Pred. No. 6.1e-66;
Matches 158; Conservative 69; Mismatches 56; Indels 10; Gaps 3;
QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASQKRTQERMSSSRPYSETIRNVISHVSKAT 60
Db 1 MAGGREIKTKISVQNTKRVTRALEMVSASKIRKAQERMKTSRPAQAMQVIGHLAQAS 60
QY 61 IGYKHPFLVDRE-VKKVGMIVVSTDRLCGGLNVNLFKTVLNMKEWKEKDVSVQLSLIG 119
Db 61 TDYQHPFLVEREQVRGVYIVISSDRGLAGLNNLFRKMLGEVRPMQDKGAEIDVVVTIG 120
QY 120 SKSINFFQSLGKILTDQSGIGDTPSVEQLIGSVNSMIDAYKKGEDVVDVYLVNKFINTM 179
Db 121 OKASAFFRIKVMVGSVTHLGDSPHIEQLVGVIKWMLDAFTGKVDVRLVYVNRFNVTM 180
QY 180 SKQPVLEKLIPIPELDNDELGERKQV---WDYIYEPDAKVLNLLVRYLESQVQAAY 236
Db 181 TQKASFQELLP-----AAEHKVAHHDWDYIYEPDAATVLEHVMTRYTESLVYQAVLE 234
QY 237 NLASEQARPMVAMKAATDNAGLNLQLVNKAQASITNELNEIVAGAAI 289
Db 235 NVASEHAARPMVAMKAASDNANKMIGTLQLVNKAQQAATQEISEIVSGAAV 287

RESULT 13
US-10-369-493-16172
; Sequence 16172, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16172
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-16172

Query Match 58.6%; Score 838; DB 14; Length 287;
Best Local Similarity 53.9%; Pred. No. 6.1e-66;
Matches 158; Conservative 69; Mismatches 56; Indels 10; Gaps 3;
QY 1 MAGAKEIRTKIASVKSTOKITKAMEMVAASQKRTQERMSSSRPYSETIRNVISHVSKAT 60
Db 1 MAGGREIKTKISVQNTKRVTRALEMVSASKIRKAQERMKTSRPAQAMQVIGHLAQAS 60
QY 61 IGYKHPFLVDRE-VKKVGMIVVSTDRLCGGLNVNLFKTVLNMKEWKEKDVSVQLSLIG 119

```

Db      61 TDYQHPFLVEREQKRGYIVISSDRGLAGLNNLFRKMLGEVRFPQKGAIDVVIIG 120
QY      120 SKSNFFQSLGKILTDGSGIGDTPSVEQLIGSVNSMIDAYKGEVDVYVLYVYKFNFTM 179
Db      121 QKASAFFRIKVMVSVTHLGDSPHIEQLGVGVKVMKDAFTGKVDVYVLYVYKFNFTM 180
QY      180 SQKPVLEKILPLPELDNDELGERKQV---WDYIYEPDAKVLNLDNLVRYLESQVYAAVE 236
Db      181 TQKASFEQLPLP-----AAEHKVAHDDNDYLYEPDAATVLEHVTWTRYIESLVYQAVLE 234
QY      237 NLASEQAAARMVAMKAATDNAGNLINELQVYVYKARQASITNELNEIVAGAAAI 289
Db      235 NVASEAARMVAMKAASDNANKMIGTLQVYVYKARQAAITQEISEIVSGAAAV 287

```

```

RESULT 14
US-10-369-493-4487
; Sequence 4487, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/369,493
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4487
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-369-493-4487

```

```

Query Match      56.9%; Score 812.5; DB 14; Length 292;
Best Local Similarity 53.4%; Pred. No. 1.1e-63;
Matches 158; Conservative 63; Mismatches 62; Indels 13; Gaps 3;

QY      1 MAGAKBIRTKIASYKSTOKITKAMEMVAASKMKTKTOERMSSSRPYSETIRNVISHVSKAT 60
Db      1 MAGWKEIRGKIKSVQNTKTKITKAMEMVAASKMKTKTOERMSSSRPYSETIRNVISHVSKAT 60
QY      61 IGYKHPFLVDRE-VKKVGMIVVSTDRGLOGGLNVLNFKTVLNEMKEWKEKOVSVQSLIG 119
Db      61 PEYRHPFMVSNEGAKTAGIILVTTDKLGGGMNTNVLRSLOKFKELGQGGKTIEATAIG 120
QY      120 SKSNFFQSLGKILTDGSGIGDTPSVEQLIGSVNSMIDAYKGEVDVYVLYVYKFNFTM 179
Db      121 TKGLGFLNLRKAVSVNVHLGDTPHLEKLGAVKQVLDLYSEGKVSAYLAYTRFVNTM 180
QY      180 SQKPVLEKILPLPELDNDELGERKQ-----WDYIYEPDAKVLNLDNLVRYLESQVY 231
Db      181 KOBEVIEQLPL-----SADQFERKEEDGTTPTQWDYIYEPDAQAVVDLIVRYVEALVY 236
QY      232 QAAVENLASQAARMVAMKAATDNAGNLINELQVYVYKARQASITNELNEIVAGAA 287
Db      237 QAAVENWASEQARMVAMKAASDNKTVINELQVYVYKARQAAITKELSEIVGGAA 292

```

```

RESULT 15
US-10-369-493-7246
; Sequence 7246, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

```

```

; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/369,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7246
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-369-493-7246

```

```

Query Match      56.9%; Score 812.5; DB 14; Length 292;
Best Local Similarity 53.4%; Pred. No. 1.1e-63;
Matches 158; Conservative 63; Mismatches 62; Indels 13; Gaps 3;

QY      1 MAGAKBIRTKIASYKSTOKITKAMEMVAASKMKTKTOERMSSSRPYSETIRNVISHVSKAT 60
Db      1 MAGWKEIRGKIKSVQNTKTKITKAMEMVAASKMKTKTOERMSSSRPYSETIRNVISHVSKAT 60
QY      61 IGYKHPFLVDRE-VKKVGMIVVSTDRGLOGGLNVLNFKTVLNEMKEWKEKOVSVQSLIG 119
Db      61 PEYRHPFMVSNEGAKTAGIILVTTDKLGGGMNTNVLRSLOKFKELGQGGKTIEATAIG 120
QY      120 SKSNFFQSLGKILTDGSGIGDTPSVEQLIGSVNSMIDAYKGEVDVYVLYVYKFNFTM 179
Db      121 TKGLGFLNLRKAVSVNVHLGDTPHLEKLGAVKQVLDLYSEGKVSAYLAYTRFVNTM 180
QY      180 SQKPVLEKILPLPELDNDELGERKQ-----WDYIYEPDAKVLNLDNLVRYLESQVY 231
Db      181 KOBEVIEQLPL-----SADQFERKEEDGTTPTQWDYIYEPDAQAVVDLIVRYVEALVY 236
QY      232 QAAVENLASQAARMVAMKAATDNAGNLINELQVYVYKARQASITNELNEIVAGAA 287
Db      237 QAAVENWASEQARMVAMKAASDNKTVINELQVYVYKARQAAITKELSEIVGGAA 292

```

Search completed: November 9, 2004, 10:00:08
Job time : 141 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 9, 2004, 05:20:38 ; Search time 23949 Seconds
(without alignments)
570.659 Million cell updates/sec

Title: US-09-545-199F-4

Perfect score: 1429

Sequence: 1 MAGAKEIRTKIASVKSTQKI.....ARQASITNELNEIVAGAAAI 289

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USFO_spool/US09545199/runat_08112004_112526_20949/app_query.fasta_1.455
-DB=genembl -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -WATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09545199 @CGN 1 1 7406 @runat_08112004_112526_20949 -NCFU=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.btg:*
3: gb.in:*
4: gb.on:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1429	100.0	870	1 AF237922	Pasteurel
2	1429	100.0	1972	6 BD252118	Anti-bact
3	1429	100.0	1972	6 AX553669	Sequence
4	1429	100.0	10229	1 AE006185	Pasteurel

C 5	1173	82.1	13828	1	U32730	U32730 Haemophilus
C 6	1173	82.1	110000	6	AR274513_04	Continuation (5 of
C 7	1173	82.1	110000	6	AR274513_05	Continuation (6 of
C 8	1132.5	79.3	867	6	BD252190	Anti-bact
C 9	1132.5	79.3	867	6	AX553798	Sequence
C 10	1121.5	78.5	304559	1	AE017151	Haemophil
C 11	1077.5	75.4	866	6	AX553832	Sequence
C 12	1004	70.3	876	6	AR378271	Sequence
C 13	1004	70.3	349483	1	BX571859	Photorhab
C 14	998.5	69.9	12928	1	AE004342	Vibrio ch
C 15	993	69.5	64019	1	EX950851_50	Continuation (51 o
C 16	991	69.3	884	6	AX188943	Sequence
C 17	991	69.3	2835	1	ECATPY	ECATPY
C 18	991	69.3	7141	1	ECOHATP	ECOHATP
C 19	991	69.3	7421	1	AF188265	Salmonell
C 20	991	69.3	11999	1	AE005605	Escherich
C 21	991	69.3	12599	1	AE015388	Shigella
C 22	991	69.3	14526	1	ECUNC	ECUNC
C 23	991	69.3	110000	1	U00096_39	Continuation (40 o
C 24	991	69.3	136254	1	ECCJW82	ECJW82
C 25	991	69.3	280900	1	AP002566	Escherich
C 26	991	69.3	293558	1	AE016991	Shigella
C 27	991	69.3	301566	1	AE016769	Escherich
C 28	989	69.2	11099	1	AE014015	Versinia
C 29	989	69.2	203728	1	AJ414160	Versinia
C 30	989	69.2	229338	1	AE017142	Versinia
C 31	988	69.1	7881	1	ECUNC	ECUNC
C 32	988	69.1	20518	1	AE008880	Salmonell
C 33	982.5	68.8	10718	1	AE015907	Shewanell
C 34	980.5	68.6	104755	1	AP005343	Vibrio vu
C 35	980.5	68.6	299938	1	AE016800	Vibrio vu
C 36	978	68.4	894	6	AR386716	Sequence
C 37	978	68.4	245050	1	AL627280	Salmonell
C 38	978	68.4	300592	1	AE016846	Salmonell
C 39	976.5	68.3	7402	1	VAUNC	VAUNC
C 40	970.5	67.9	288108	1	AP005083	Vibrio pa
C 41	947.5	66.3	251097	1	CR378674	Photobact
C 42	931.5	65.2	7863	1	AB035129	Colwellia
C 43	926.5	64.8	16662	1	AE004967	Pseudomon
C 44	915	64.0	349544	1	CR378675	Photobact
C 45	896.5	62.7	142665	1	AE016794	Pseudomon

ALIGNMENTS

RESULT 1	AF237922	Pasteurella multocida ATP synthase F1 gamma chain (atpg) gene, complete cds.	870 bp	DNA	linear	ECT 08-MAY-2000
LOCUS	AF237922	Pasteurella multocida ATP synthase F1 gamma chain (atpg) gene, complete cds.				
ACCESSION	AF237922					
VERSION	AF237922.1	GI:7716504				
KEYWORDS						
SOURCE		Pasteurella multocida				
ORGANISM		Pasteurella multocida				
REFERENCE		Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Pasteurella.				
TITLE		1 (bases 1 to 870)				
AUTHORS		Fuller,T.E., Kennedy,M.J. and Lowery,D.E.				
DATE		Identification of Pasteurella multocida virulence genes in a septicemic mouse model using signature-tagged mutagenesis				
JOURNAL		Unpublished				
REFERENCE		2 (bases 1 to 870)				
AUTHORS		Fuller,T.E., Kennedy,M.J. and Lowery,D.E.				
TITLE		Direct Submission				
JOURNAL		Submitted (24-FEB-2000) Discovery Research, Pharmacia & Upjohn Animal Health, 7923-25-434, 7000 Portage Road, Kalamazoo, MI 49001-0199, USA				
FEATURES		Location/Qualifiers				
source		1..870				
		/organism="Pasteurella multocida"				
		/mol_type="genomic DNA"				
		/db_xref="taxon:747"				

```

gene      1..870
          /gene="atpg"
CDS       1..870
          /gene="atpg"
          /codon_start=1
          /transl_table=11
          /product="ATP synthase F1 gamma chain"
          /protein_id="AAF8408.1"
          /db_xref="GI:7715505"
          /translation="MAGAKETIKIASVSTQKITKAMENVAASKVKRKTQERMSSRP
YSETRIVSHVSKATIGYHPFLDREVKKGMIVVSTDRGLCGGLNVNLEKTVLNE
MKWEKDVSVQLSLIGSKINPFSQIGIKILQDSIGIDTPSEVLGICVSNMIDAY
KKGEDVYVLYNFKPINTMSQKPLEKLIPLPELDNDELGERKQVWDYLYEPDAKVL
DNLVLYLESQVQAAVENLASQAARFWMKAAATDNAGNLINELQLVYNKARQASIT
NELNEIVAGAAI"

ORIGIN
Alignment Scores:
Pred. No.: 1,48-110 Length: 870
Score: 1429.00 Matches: 289
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-545-199F-4 (1-289) x AF237922 (1-870)
QY 1 MetAlaGlyAlaLysGluLeuAArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
DB 1 ATGCAGCGTCTAAAGAGATAAGACCAAAATCGCAGTGTAAAGATACACAAAAAATT 60
QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
DB 61 ACTAAAGCGATGAAATGGTGTGCTCGAATCGTAAACGCAAGCCATGTCT 120
QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
DB 121 TCTTCAGCCCTTATTCAGAAACAATACGTACGTGATTAGCCGTTTCCAAAGCAACG 180
QY 61 IleGlyTyrLysHisProPheLeuValAlaAspArgGluValLysValGlyMetIleVal 80
DB 181 ATTGGTTTACAAGCATCCATTTTGTAGTGATCGCAAGTAAAAAAGTGGCATGATGTT 240
QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLysPheLysThrValLeu 100
DB 241 GTGTCCACAGATCGTGCTTTGTGTGCTTTAAACGTGAATCTGTTTAAACATGTTATTA 300
QY 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
DB 301 AATGAAATGAAAGATGGAAGAAAAAGATGTTCCGTTCAATTGAGTTTAAATCGGTCT 360
QY 121 LysSerIleAsnPhePheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
DB 361 AAATCTATCACTTTTCCAACTTTGGGAATTTAAATTTTAAACCAAGATTCAGGTATT 420
QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
DB 421 GGTGATACCTCCCTCTGTCAGCAGTAAATTTGTTCCAGTCAATTCATGATTGATGCTTAT 480
QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
DB 481 AAAAAGGGGAAGTAGATGTTGTGATTAGTTTATATATAACAAATTTTAAACAGATGTCG 540
QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
DB 541 CAAAGCCAGTATTGAAAAATTAATTCATTAACAGATTTAGATTAATGATGATTAGGC 600
QY 201 GluArgLysGlnValTyrAspTyrIleTyrGlnProAspAlaLysValLeuLeuAspAsn 220
DB 601 GAAAGAAACAAGATTGGGATTATATTACGAACCTGATGCGAAAGTATTATTAGATAAT 660
QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluLeuLeuAlaSer 240

```

```

Db 661 TTATTGGTTCGTTATTAGAAATCTCAGGTTTATCAAGCAGCAGTTGAAACCTTCTTCT 720
QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
DB 721 GAGCAAGCCCTCGAATGGTTCGCCATGAAAGCAGCAACAGATACGAGGTAACTTAATT 780
QY 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
DB 781 AATGAGTTACAGTTAGTCTATACAAAGCTCGTCAAGCAAGTATTACAAATGAATTAAT 840
QY 281 GluIleValAlaGlyAlaAlaIle 289
DB 841 GAAATGTTCCGCGTCAGAGCAAT 867

RESULT 2
BD252118 1972 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Anti-bacterial vaccine compositions.
BD252118
ACCESSION
VERSION
BD252118.1 GI:33061888
KEYWORDS JP 2002541790-A/2.
SOURCE Pasteurella multocida
ORGANISM Pasteurella multocida
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Pasteurella.
REFERENCE
1 (bases 1 to 1972)
LOWERY,D.E., FULLER,T.E. and Kennedy,M.C.
Anti-bacterial vaccine compositions
Patent: JP 2002541790-A 2 10-DEC-2002;
PHARMACIA AND UPJOHN CO
COMMENT
OS Pasteurella multocida
PN JP 2002541790-A/2
PD 10-DEC-2002
PP 06-APR-2000 JP 2000611649
PR 09-APR-1999 US 60/128689,10-SEP-1999 US 60/153453 PI
DAVID E LOWERY, TROY E FULLER, MICHAEL J KENNEDY PC
C12N15/09,A61K39/02,A61K39/102,A61K39/395,A61P31/04,C07K14/285, PC
C07K16/12,
PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/01,C12P21/02 PC
C12P21/08,C12Q1/18,
PC G01N33/15,G01N33/50,G01N33/53,G01N33/569,G01N33/577,C12N15/00,
PC C12N5/00,
PC C12N5/00
CC atpg
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1972
/organism="Pasteurella multocida"
/db_xref="taxon:747"

ORIGIN
Alignment Scores:
Pred. No.: 3,85e-110 Length: 1972
Score: 1429.00 Matches: 289
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-545-199F-4 (1-289) x BD252118 (1-1972)
QY 1 MetAlaGlyAlaLysGluLeuAArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
DB 364 ATGCAGCGTCTAAAGAGATAAGACCAAAATCGCAGTGTAAAGATACACAAAAAATT 423
QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
DB 424 ACTAAAGCGATGGAATGGTGTGCTCGAATCGTAAACGCAAGCCATGTCT 483
QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60

```

```
Db 484 TCTTCAGCGCCTTATTTCAGAAACAATACCTAACGCTGATTAGCCACGTTTCCAAAGCAACG 543
Qy 61 ILeGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
Db 544 ATTGGTTTACAAGCATCCATTTTATTAGTGATCGCGAAGTAAATAAAGTAGTGGCATGATTGTT 603
Qy 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db 604 GTCTCACAGATCGTGGTCTTTGTGGTGGCTTAAACGTAACCTGTTTAAACATGTTATTA 663
Qy 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnSerLeuIleGlySer 120
Db 664 AATGAATGAAGATGGAAGAAAGATGTTTCGGTTCATTCAGTTTAAATCGGTCT 723
Qy 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db 724 AAATCTATCAACTTTTTCATCTTTGGGAATTAATAATTTTAAACCAAGATTCAGGTATT 783
Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db 784 GGTGATACCTCCCTCTGTGTAGCAGTAAATTTAGTGTTCAGTCAATTCATGATTCGCTTAT 843
Qy 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 844 AAAAAGGGGGAAGTAGATCTTTGGAATTAATAATTTTAAACCAAGATTCAGGTATT 783
Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db 784 GGTGATACCTCCCTCTGTGTAGCAGTAAATTTAGTGTTCAGTCAATTCATGATTCGCTTAT 843
Qy 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 844 AAAAAGGGGGAAGTAGATCTTTGGAATTAATAATTTTAAACCAAGATTCAGGTATT 783
Qy 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db 904 CAAAGCCAGTATTTGGAAATAATTAATTCATTCACGAAATTAAGTATGATGAATTAGGC 963
Qy 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db 964 GAAAGAAAACAAGTTTGGGATATATTTACGAACTGATGCGAAAGTATTATTAGATAAT 1023
Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaLysValLeuLeuAspAsn 240
Db 1024 TTATGTGTTTGGTATTTAGAACTCTCAGGTTTATCAAGCAGCAGTGTGAAACCTTGCTTCT 1083
Qy 241 GluGlnAlaAlaArgMetValAlaMetLysAlaLysValAsnValAsnLeuIle 260
Db 1084 GAGCAAGCGCTCGAATGTCGCCATGAAGCAGCAACAGATTAACGAGGTAACTTAAT 1143
Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db 1144 AATGAGTTACAGTTAGTCTTATTAACAAGCTCGTCAAGCAAGTATTACAAATGAATTAAT 1203
Qy 281 GluIleValAlaClyAlaAlaIle 289
Db 1204 GAAATGTTGCGGTGCGAGCAAT 1230
RESULT 3
AX553669 LOCUS 1972 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 3 from Patent WO02075507.
ACCESSION AX553669
VERSION AX553669.1 GI:25897667
KEYWORDS Pasteurella multocida
SOURCE Pasteurella multocida
ORGANISM Pasteurella multocida
REFERENCE 1
AUTHORS Lowery D.E., Fuller T.E. and Kennedy M.J.
TITLE Anti-bacterial vaccine compositions
JOURNAL Patent: WO 02075507-A 3 26-SEP-2002;
Pharmacia & Upjohn Company (US)
FEATURES
source
1. .1972
/organism="Pasteurella multocida"
/mol_type="unassigned DNA"
/db_xref="taxon:747"
364. .1233
/gene="atpG"
```

CDS

```
364. .1233
/gene="atpG"
/codon_start=1
/transl_table=11
/protein_id="CAD58126.1"
/db_xref="GI:25897668"
/translation="MAGAKIRTKIAVSKSTQKITKAVEMVAASKKRKTKQRMSSSRP
YSETRKNVISHVSKATIGKHPFLVDREVKKVGMIVVSTDRGLCGGLNVNLFKTVLNE
MKEMKEDVSVQLASIGSKSINFFQSLGKILITQDSIGIGTSPVEQLIGSVNSXIDAY
KKGVDVVLVLYNKFINTMSKPLVKLILFLDELDELGERKQWDYIYEPAKVLL
DNLVRYLESQYQRAVENLASEQARVAVKAAATDAGNLINELOLVYNKARQASIT
NELNEIVAGAAAI"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 3,85e-110 Length: 1972
Score: 1429.00 Matches: 289
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
```

US-09-545-199F-4 (1-289) x AX553669 (1-1972)

```
Qy 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
Db 364 ATGGCAGGTGCTAAAGAGATAAGAACCAAAATCCGAGTGTAAATAAGTACACAAAAAATT 423
Qy 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
Db 424 ACTAAAGCGATGGAATGTTGCTCGCTCGAAATGCTAAACGCGAAGACGATGCT 483
Qy 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db 484 TCTTCACGCCCTTATTACAAACAATACGTAACGTCGATTCAGCCACGTTTCCAAAGCAACG 543
Qy 61 ILeGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
Db 544 ATTGGTTTACAAGCATCCATTTTATTAGTGGATCGGAAGTAAATAAAGTAGTGGCATGATTGTT 603
Qy 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db 604 GTGTCCACAGATCGTGGTCTTTGTGGTGGCTTAAACGTAACCTGTTTAAACATGTTATTA 663
Qy 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnSerLeuIleGlySer 120
Db 664 AATGAATGAAGATGGAAGAAAGATGTTTCGGTTCATTCAGTTTAAATCGGTCT 723
Qy 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db 724 AAATCTATCAACTTTTTCATCTTTGGGAATTAATAATTTTAAACCAAGATTCAGGTATT 783
Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db 784 GGTGATACCTCCCTCTGTGTAGCAGTAAATTTAGTGTTCAGTCAATTCATGATTCGCTTAT 843
Qy 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 844 AAAAAGGGGGAAGTAGATCTTTGGAATTAATAATTTTAAACCAAGATTCAGGTATT 783
Qy 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db 904 CAAAGCCAGTATTTGGAAATAATTAATTCATTCACGAAATTAAGTATGATGAATTAGGC 963
Qy 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db 964 GAAAGAAAACAAGTTTGGGATATATTTACGAACTGATGCGAAAGTATTATTAGATAAT 1023
Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaLysValLeuLeuAspAsn 240
Db 1024 TTATGTGTTTGGTATTTAGAACTCTCAGGTTTATCAAGCAGCAGTGTGAAACCTTGCTTCT 1083
Qy 241 GluGlnAlaAlaArgMetValAlaMetLysAlaLysValAsnValAsnLeuIle 260
```

```

Db      1084 GAGCAAGCGCTCGAATGTCGCCATGAAGCAGCAACAGATAACGACGTAACCTTAATT 1143
QY      261 AenGluLeuGlnLeuValTyAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db      1144 AATGAGTTACAGTATAGTTATTAACAAGCTGCTACAGCAAGTATTACAAATGAATTAAT 1203
QY      281 GlnLeuValAlaGlyAlaAlaLalle 289
Db      1204 GAAATGTTGCGGTGCAGCACAATT 1230

RESULT 4
LOCUS   AE006185                10229 bp    DNA        linear    BCT 16-JUN-2004
DEFINITION Pasteurella multocida subsp. multocida str. Pm70 section 152 of 204
of the complete genome.
ACCESSION AE006185 AE004439
VERSION   AE006185.1 GI:12721868
KEYWORDS
SOURCE   Pasteurella multocida subsp. multocida str. Pm70
ORGANISM Pasteurella multocida subsp. multocida str. Pm70
          Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
          Pasteurellaceae; Pasteurella.
REFERENCE 1 (bases 1 to 10229)
AUTHORS   May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whittam, T.S. and
          Kapur, V.
TITLE     Complete genomic sequence of Pasteurella multocida, Pm70
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
MEDLINE   21145866
PubMed   11248100
REFERENCE 2 (bases 1 to 10229)
AUTHORS   Zhang, Q. and Kapur, V.
TITLE     Direct Submission
JOURNAL   Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
          University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
          55108, USA

FEATURES             Location/Qualifiers
     1..10229
     /organism="Pasteurella multocida subsp. multocida str.
     Pm70"
     /mol_type="genomic DNA"
     /strain="PM70"
     /sub_species="multocida"
     /db_xref="taxon:272843"
     275..2164
     /gene="gidA"
     /note="synonym: PM1485"
     275..2164
     /gene="gidA"
     /codon_start=1
     /transl_table=11
     /product="GidA"
     /protein_id="AAK03569.1"
     /db_xref="GI:12721869"
     /translation="MFYTENYDVI VIGGSHAGTEAALAPAWGLKTLITHTVDTLQ
     MSCNPAIGIGIKGLHVRIDEADGGLMATDAQAGIQFTLSSKGPVAVRATRAQDRV
     LVQAVRIALENCENLDIQOEVTBIIDQRCVGVVTMGLKFKHAKAVILTAGTFLS
     GKIHIGLNTYAGAGPASMADRLRELNLRLVRLKTPPTPRIDARTIDFSLAKQ
     HGDEKLVPFSGMSVDQHPQICPTITHTNEQTHEVIRNNLDSPMTAGIIEGIGPRY
     CPSIEDKVRFSERNSHQIYLEPGLTSENIYPNGISTSLPFDVQMKIVMSGMGEKA
     RIIKGYAIIEYFDPRDLKPTLETKSIRGLFAGQINGTTGYEEAAGQGLLAGING
     LFVQEKAAWPPRRDQAYIGVLVDLCTLTGKEPVRVPTSRAEYELLIRENADSRITP
     IAHOLGIDIEKWARFNQKMWENIELEORLQIMHRSRSLDSEANKVLSGFLVRAS
     GEDLLRPENYQILTSLTPFPAMDSEAVEQVEIATIKQYIEHQOESEIAQKRE
     STAPFAFDYTVVSGLSNEVRKAEQHRPVSIGQASRISGVTPAAISILIVSLKQGM
     LRGE"
     2164..2796
     /gene="gidB"
     /note="synonym: PM1486"
     2164..2796
     /gene="gidB"
     /codon_start=1
     /transl_table=11

```

```

/product="GidB"
/db_xref="GI:12721870"
/translation="MTNLEQQLSQKLEILLKOTALSITDQOEOLIKLVLLINKKKA
YNTSVRDPDMELVKHILDSVVSPHLOGKHRIIDVGTGPGLPGLPLAIWPNKHVLL
DSLGRISFTRNAIRBELGLDNVEAVLSRVEEYHPQPPDGVLSRAFASLKDNTWCQH
LPKQDGYFVALKGLYHQEEVEELSEKFTIQQVIRLQVPELIGERHLVIVK"
2934..3341
/gene="PM1487"
2934..3341
/gene="PM1487"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03571.1"
/db_xref="GI:12721871"
/translation="MSKVIYQAKQKYKIALGIELGVLLASLTCLIGYQWQVAISFILG
SSAIFPHCLFVNFVEFTKKQPSNKLTFYGGEAVKFTSTIFLIVLAFKLPNSMHI
AFFVGYFITVLNNILFFFTISKYLNKHPHYKQ"
3352..4146
/gene="atpB"
/note="synonym: PM1488"
3352..4146
/gene="atpB"
/codon_start=1
/transl_table=11
/product="AtpB"
/protein_id="AAK03572.1"
/db_xref="GI:12721872"
/translation="MAAELTAGYIGHLAFIKTGDSPWVHVDLTLPSTISGAIFLF
VSKVAKATPGVSKMOCFTVEIWDWIDGIVKGNFHPHRAVGLPALTIFCWYFINN
ADLPVDFPQLAHLFGLEYLRVAPTADISGTGLSGISGVFLIFFTIKSGMSGFV
KSTYHPNHLPLIPNLALSESVTLLAKPVSIAFLFCNMYAGELIFTLIAVMYMANN
PALNSGIFPMHLAWAIFHILVTLOAFIFMMLTVVYLSMGYNKAH"
4203..4457
/gene="atpE"
/note="synonym: PM1489"
4203..4457
/gene="atpE"
/codon_start=1
/transl_table=11
/product="AtpE"
/protein_id="AAK03573.1"
/db_xref="GI:12721873"
/translation="METVITITIASAILAAALGATGALFAILGGRFLESSARQPEL
ASSLQIKMFIAGLLDAISMTAVGALLFIFANPFIDLK"
4537..5007
/gene="atpF"
/note="synonym: PM1490"
4537..5007
/gene="atpF"
/codon_start=1
/transl_table=11
/product="AtpF"
/protein_id="AAK03574.1"
/db_xref="GI:12721874"
/translation="MNLNATLIGLIAFAIFVACMKFVWPPIKAIEBERQSIANAL
ASAAKQEQADTALVEQELTEKMAQAOIILANKRNEIEELVEKVEAEATKAKII
EQGVAEAEKRVQBELRVKVASLATAGAEKIVGRTVDEANSIDIDKLVAEL"
5021..5554
/gene="atpH"
/note="synonym: PM1491"
5021..5554
/gene="atpH"
/codon_start=1
/transl_table=11
/product="AtpH"
/protein_id="AAK03575.1"
/db_xref="GI:12721875"
/translation="MSBLTIAERPYPAAKAFDAIEQNAVEKWNMLHFSSELIKDET
QTLKSSLSAKLADTVISICGOLDQYQGNLIRLMAENKRLVLPALYQAFQHVVEE
HOKVAEQVISAQPLNATQKIASAMEKRLARKVKLNCSLDSSLIAGVIITDDFVI
DGSSRGQLSRLANELQL"

```

Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0
US-09-545-1998-4 (1-289) x AS006185 (1-10229)			
Qy	1	MetAlaGlyAlaLysGluLeuLeuArgThrLysLeuLeuSerValLysSerThrGlnLysLeu	20
Db	7135	ATGCGAGGTGCTAAAGAGATAGAACCAAAATCGGAGTGTAAAGAGTACACAAAATTT	7194
Qy	21	ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer	40
Db	7195	ACTAAAGCGATGGAAATGGTTGCTCCCTCGAAATGCGTAAAGCGCAAGAACCATGTCT	7254
Qy	41	SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr	60
Db	7255	TCTTCACGCCCTTATTCAGAACCAATACGTAACTGGATTAGCCACGCTTCCAAAGCAACG	7314
Qy	61	IleGlyTyrLysHisProPheIleuValAspArgLysValLysValGlyMetIleVal	80
Db	7315	ATTGGTTACAGCATCCCAATTTTATGTGGATCGCAAGTAAAGGTCGTCATGATGTT	7374
Qy	81	ValSerThrAspArgGlyLeuCysGlyLeuAsnValAsnLeuPheLysThrValLeu	100
Db	7375	GTGTCACAGATCGGTGCTTTTGTGGCTTAACTGTGACCTGACCTGTTTAAACCTGATTA	7434
Qy	101	AsnGluMetLysGluThrPylsGluLysAspValSerValGlnLeuSerLeuIleGlySer	120
Db	7435	AATGAATGAAGAGATGGAAGAAAGAAAGATGTTTCTGTTCAATTTGAGTTTAAATCGGTCT	7494
Qy	121	LysSerIleAsnPheGlnSerLeuGlyLysLeuLysLysLeuThrGlnAspSerGlyIle	140
Db	7495	AAATCTATCACTTTTCCATCTTTGGGAATTAATAATTTTAAACCAAGATTCCGATTT	7554
Qy	141	GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr	160
Db	7555	GGTGATCTCCCTCTGTTGAGCAGTTAATTTGGTTCAGTCAATTTCTATGATTGATGCTAT	7614
Qy	161	LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer	180
Db	7615	AAAAAGGGGAAGTAGATTGTTGTTATTTAGTTTATTAACAAAATTTTATTAACAGATGTCG	7674
Qy	181	GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly	200
Db	7675	CAAAAGCCAGTATTGGAAAAATTAATTTCCATTACCAGATTAGATATGATGATGATAGGC	7734
Qy	201	GluArgLysGlnValThrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn	220
Db	7735	GAAGAAAAACAAGTTTGGGATTATTTACGAACCTGATCGCAAGATATTATTAGATATAT	7794
Qy	221	LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluAsnLeuAlaSer	240
Db	7795	TTATTGGTTGTTATTAGATCTCAGGTTTATCAAGCGCAGTGTGCAAAACCTTCCTCTCT	7854
Qy	241	GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle	260
Db	7855	GAGCAAGCCGCTCGAATGGTCCCATGAAAGCAGCAACAGATAACGAGTTAACTTAAT	7914
Qy	261	AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGlnLeuAsn	280
Db	7915	AATGAGTTACAGTAGTCTATTAACAAGCTCGTCAAGCAAGTATTACAAATGAATTAAT	7974
Qy	281	GluLeuValAlaGlyAlaAlaAlaIle	289
Db	7975	GAATTTGTCGGTGCAGCAGCAAT	8001

```

Alignment Scores:
Pred. No.: 2,95e-109
Score: 1429.00
Length: 10229
Matches: 289
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
/translation="MSVNLNLTVVSAPQQIFSSQVESIQATCIQEGELGILAGHPTPLAT
IKPGIVKLTLEDGKEEVIYVSGGLEVFQVNVVYVLADTAIRGDELQDRLIAAKKRAE

```

RESULT 5	U32730/c	U32730	13828 bp	DNA	linear	BCT 02-JUN-2004
LOCUS		U32730	13828 bp	DNA	linear	BCT 02-JUN-2004
DEFINITION		Haemophilus influenzae Rd KW20 section 45 of 183 of the complete				
ACCESSION		U32730				
VERSION		U32730.1				
KEYWORDS		GI:3212191				

SOURCE
ORGANISM
REFERENCE
AUTHORS

Haemophilus influenzae Rd KW20
Haemophilus influenzae Rd KW20
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Haemophilus.

1 (bases 1 to 13828)
Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A.,
Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A.,
Murray, J.M., McKenney, K., Sutton, G.G., FitzHugh, W., Fields, C.A.,
Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A.,
Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E.,
Cotton, M.D., Utterback, T., Hanna, M.C., Spriggs, D.T., Saudek, D.M.,
Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Saudek, D.M.,
Geoghagen, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M.,
Smith, H.O. and Venter, J.C.

Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd KW20
Science 269 (5223), 496-512 (1995)
95350630
PUBMED
7542800

2 (bases 1 to 13828)
Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S.,
Borodovsky, M., Rudd, K.E. and Koonin, E.V.
Metabolism and evolution of Haemophilus influenzae deduced from a
whole-genome comparison with Escherichia coli
Curr. Biol. 6 (3), 279-291 (1996)
96398784
PUBMED
8805245

3 (bases 1 to 13828)
White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
Direct Submission
TITLE
Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

4 (bases 1 to 13828)
White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
Direct Submission
TITLE
Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

5 (bases 1 to 13828)
White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D.,
Peterson, J., Hickey, E., Dodson, R. and Winn, M.
Direct Submission
TITLE
Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

REMARK
COMMENT
FEATURES

The whole genome was shifted by 588 nucleotides for a new start
On Jun 12, 1998 this sequence version replaced gi:1573448.
Location/Qualifiers

1..13828
/organism="Haemophilus influenzae Rd KW20"
/mol_type="genomic DNA"
/db_xref="taxon:71421"
65..1168
/gene="HI0470"
65..1168
/gene="HI0470"
/note="similar to SP:P06986 GB:X03416 PID:41695
PID:41710 percent identity: 60.40; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="histidinol-phosphate aminotransferase (hisC)"
/protein_id="AAC22129.1"
/db_xref="GI:1573449"

/translation="MNDYNSHILQWKKTMTITLSTRONIQALTPYQSARKLGNGTII
WLNANEYTSPEFQSLKDLNRYPEPQPRQVQVANYAGVSTENVLVTRGGDEGIEL
IHTFCBQKQAILFCPTPTGMVAVSAETAGVSKVPLTDDPOLNPEIKHNDVK
VVFVCSNNPTGNLLKQSDLDLQILQITAGKAVIVVDEAVIEPCPSVNLKNTPHL
AIITLSKAPALAGLRCFGLVLANPELIDILSKVIAPYPIPVPSADLAELQALRPSNIAT

VOALTOELLSNRCWLAKALLLVHQVEKYVSEANVLLIKQNGOAVFKALWEQGIILR
DNQKLHLQNCRIIVTGRNECEKVVEAIKEVK"
1280..2368
/gene="HI0471"
1280..2368
/gene="HI0471"
/note="similar to SP:P06987 GB:X03416 PID:41696 PID:41711
GB:U00096 percent identity: 64.97; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="amidazoleylycerol-phosphate dehydratase /
histidinol-phosphatase (hisB)"
/protein_id="AAC22130.1"
/db_xref="GI:1573450"
/translation="MQPTLFTDRDGLIDEPKTDFOIDSLKLEPKVPIALLRLKA
KYRFVIVSNQDGLGTDAFPQDFDPKPHVMALFESQGITFDEVILCPKPENCICLR
KPKILLDHYIRKNLFDIDRSFVIGDRETDVQLAENLGRIRAIQYDPQKMDLIAEKL
LGRTVNCGRKPPRFAEVIQRTKEDIKVQVLDGAVNEIKTGVGFDMQLQIATH
GGFRNVQCKGDLWIDEHTVDTALGQALQKQAVGDKRGIRAFPGVFLPMDECKABC
ALDLSGRPWIKFNACFKRDKVGFSTELTEHFFQSLAFSLATLHLNVTGNDHDKIE
SLFKAFGRTLRQAIRIEGNEPSSKGLV"
2434..3033
/gene="HI0472"
2434..3033
/gene="HI0472"
/note="similar to SP:P10375 PID:41712 GB:U00096
PID:1736701 PID:1736710 percent identity: 55.90;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="amidotransferase (hisH)"
/protein_id="AAC22131.1"
/db_xref="GI:1573451"
/translation="MINTITDTCANLSVKEAFDRLYGNTBITFDLNKIKSADKLI
LFGPTGTAAMYLQERLIETQNTQVPLGICLGMQMTSESGNVTPLNLISGK
TNRIPTGLPQPMQWNRVQFVKNCLPFDGIVQNSHFYFVHSYAVSFNEHSVAISNG
VNEFAAKENFGVQVQHPERSGKNGALLKNFVKVFF"
3069..3818
/gene="HI0473"
3069..3818
/gene="HI0473"
/note="similar to SP:P10371 PID:41713 GB:U00096
PID:1736702 PID:1736711 percent identity: 62.86;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphoribosylformimino-5-aminimidazole
carboxamide ribotide isomerase (hisA)"
/protein_id="AAC22132.1"
/db_xref="GI:1573452"
/translation="MKQSIIPALDLINGVVRHQDYAKQTYSDNPIKQFDNVR
QGAQLHLVDLTGAKNPQSRQALIGKIYEATQCKVQVGGRTBCDVAIDLAVGNR
VWIGSTAVTRSHVQWVFKYGAKEFVLADVNINASGQKIVAIQSGWEEGVLLETL
IEDPTVGLQQLVCTDISRDGTLTSGNIYQIECKYPIQFQSSGGIGSLADIEAL
KEGTGSGVIVRALLGKFTLSEAICWQNG"
3800..4576
/gene="HI0474"
3800..4576
/gene="HI0474"
/note="similar to SP:P10373 PID:41714 PID:598474 GB:U00096
PID:1736704 percent identity: 82.03; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hisF cyclase (hisF)"
/protein_id="AAC22133.1"
/db_xref="GI:1573453"
/translation="MLAKXELFCLDVRQGVKQVQFRNHEIIGDIVPLAQRVQEGA
DELVFDITASSDGRVTDSKSWERVAQVIDIPFCVAGGKITIEDAKLIFAGDKI SI
NSPALADPILRLADRFQVQAIIVGIDSWFEQETKYWNQYVCTDETRTRQWMLL
DWFVQOCCAGEIIVLNMNODGLRGYDLAQLKLVRCVCRVPLIASGAGKMWHRD
AFTEAKVDGALAAVPHKQIIEIGELKSYLVQSAIETRSE"

gene
CDS

gene
CDS

gene
CDS

```

gene      4576..5241
CDS       /gene="HI0475"
          4576..5241
          /gene="HI0475"
          /note="similar to SP:P06989 GB:D43637 GB:U02072 GB:X03974
          PID:41700 percent identity: 61.22; identified by sequence
          similarity/putative"
          /codon_start=1
          /transl_table=11
          /product="phosphoribosyl-AMP cyclohydrolase /
          phosphoribosyl-ATP pyrophosphohydrolase (hisIE)"
          /protein_id="AAC22134.1"
          /db_xref="GI:1573454"
          /translation="MNITKIDWQKNGLLPVIQNIISTREVLMGLMNEEALTKTIKE
          KVTFFPRTKQRLWKGISGNFENVEMSLOCNDLILVDPIGATGTCHEYSCFH
          QFTSPQENKQGFANWAFIKLEQHLKKNKADPENSYTALHAKGTKKIAQKVGEE
          GVETLAAVAQDAKAEVISELDLVYHLTVLLHNDQLQWYEIIIAKLQHQHIGLHPG
          GNK"
gene      5323..5475
CDS       /gene="HI0476"
          5323..5475
          /gene="HI0476"
          /note="hypothetical protein; identified by Genemark;
          putative"
          /codon_start=1
          /transl_table=11
          /product="H. influenzae predicted coding region HI0476"
          /protein_id="AAC22144.1"
          /db_xref="GI:1573464"
          /translations="VWIKCIDKQONLGNILFLLLKQYSKEDSKFTIYKFLQTVN
          YTIQLS"
gene      5505..6707
CDS       /gene="HI0477"
          5505..6707
          /gene="HI0477"
          /note="similar to GB:M23240 SP:P18199 PID:148089 GB:U00096
          PID:1736569 percent identity: 47.24; identified by
          sequence similarity; putative"
          /codon_start=1
          /transl_table=11
          /product="tyrosine-specific transport protein (tyrP)"
          /protein_id="AAC22135.1"
          /db_xref="GI:1573455"
          /translations="MNKTVGSLLVAGNMGKMLAMPLTSAGIFGFTLVLLGLWA
          LUTFSALLVEVLYQASDAGIQAQYFGKTRIIATAVLIIFLVAALAAVLSGGG
          SLKDLPSFGDKSVSLFTVIFGSPVIFGHSVDKINRVLFVFLAAFAVLSML
          PEIKFDNLMATPIDKALIIISAPVPTAFGPHGSPSLNKYLDGNVKALRFSILVGA
          ITLCAYILWQLSTHGLLTQNEFLQILKEDATLNGLVKATFAITGSGNVIASVAKLFSTL

Alignment Scores:
Pred. No.: 1,le-87      Length: 13828
Score:      1173.00      Matches: 221
Percent Similarity: 91.00%      Conservative: 42
Best Local Similarity: 76.47%      Mismatches: 26
Query Match: 82.09%      Indels: 0
DB:      1      Gaps: 0

US-09-545-199F-4 (1-289) x U32730 (1-13828)

QY      1 MetAlaGlyAlaLysGluLeuArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
Db      9610 ATGGCAGGTGCAAAAGAGATAAAACCAAAATTGCCAGTGACAAAGTACACAAATAATC 9551
QY      21 ThrLysAlaMetGluMetValAlaLaserLysMetArgLysThrGlnGluArgMetSer 40
Db      9550 ACTAAGCAATGGAATGTGGCAACCTCGAAATTCGTAACCAACGAGGATCGATGGCT 9491
QY      41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db      9490 GCATCTCGTCGGTATTCTGAAACTATCGTACCGTTATTACTCATGTGCTAGGCAAGT 9431
QY      61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80

```

```

Db      9430 ATCGGTTATAAACAATCGTTCTTAGTTGAGCGCGAAGTGAAGAAAATCGGTATCTGGTT 9371
QY      81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db      9370 ATTTCACACAGATCGTGGGATGTGTGGTGGTTAAATGTTAAATTTATTTCAAAACACACTT 9311
QY      101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
Db      9310 AACCAATAAAAAATTGGAAAGAACAAATATTCTACAGATTGGGCTTAATAGTTCA 9251
QY      121 LysSerIleAsnPhePheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db      9250 AAAGGGATTAGTTTTCGGTTCCTTTGGATTAAATATCAAAAGTCAGCTTCTGCTTTA 9191
QY      141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db      9190 GCGGATAGCCGCTCTAGAGAAATTAATGTGTGGCAATACAAATGTTTCTGCTTAT 9131
QY      161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db      9130 CGTAATGGTGAATTCATGCAGTTTATATTCATACATAAATTTGTTAATACGATGTCG 9071
QY      181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db      9070 CAAAAGCCTGTGTACACAAATAGTTCCTTTACAGAAATCTAAAGACGATCAATTAAT 9011
QY      201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db      9010 GAAAGACAAACAGACTTGGGATTATCTTTATGAGCCAGAACCAAAAGTAGTACTATTAGATAGC 8951
QY      221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluAsnLeuAlaSer 240
Db      8950 CTTTATAGTTCGTTATTAGTCCCAAAATTTATCAAGCGGTTGTAGATAATGTAGCTTCA 8891
QY      241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
Db      8890 GAACAGCGGCTCGAATGCTAGCAATGAACAGCAACCGAATATGCAAGTAAATTTAATT 8831
QY      261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db      8830 AATGATCTCGGTTGGTGATCAACAAACCTGCTCAAGCAAGTATCACAAATGAATGAAT 8771
QY      281 GluIleValAlaGlyAlaAlaAlaIle 289
Db      8770 GAATCGTAGCGGTCGCGCGCGGATT 8744

RESULT 6
AR274513_04/c
WPCOMMENT
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513
Fragment Name      Begin      End
AR274513_00      1      110000
AR274513_01      100001    210000
AR274513_02      200001    310000
AR274513_03      300001    410000
AR274513_04      400001    510000
AR274513_05      500001    610000
AR274513_06      600001    710000
AR274513_07      700001    810000
AR274513_08      800001    910000
AR274513_09      900001   1010000
AR274513_10     1000001   1110000
AR274513_11     1100001   1210000
AR274513_12     1200001   1310000
AR274513_13     1300001   1410000
AR274513_14     1400001   1510000
AR274513_15     1500001   1610000
AR274513_16     1600001   1710000
AR274513_17     1700001   1810000
AR274513_18     1800001   1930121

Continuation (5 of 19) of AR274513 from base 400001 (AR274513 Sequence 1 from patent US
Alignment Scores:

```

Pred. No.: 1,42e-86 Length: 110000
Score: 1173.00 Matches: 221
Percent Similarity: 91.00% Conservative: 42
Best Local Similarity: 76.47% Mismatches: 26
Query Match: 82.09% Indels: 0
DB: 6 Gaps: 0

US-09-545-199F-4 (1-289) x AR274513_04 (1-110000)

QY 1 MetAlaGlyAlaLysGluThrLysLeuValSerValLysSerThrGlnLysLeu 20
DB 101934 ATGCGAGTGCCTGCAAGAGATACCAAAATTCGAGTACAGATACCAAAATC 101875

QY 21 ThrLysAlaMetGluMetValAlaLysSerValLysSerThrGlnLysMetSer 40
DB 101874 ACTAAGCAATGCAATGCGCAACCTGCAAAATCGTAAACGCAAGATCGTATGCT 101815

QY 41 SerSerArgProTyrSerGluThrLysLeuValSerValLysSerThrGlnLysAlaThr 60
DB 101814 GCATCTCGTCGATTCGAACTATCGTACGTTATGATGATGCTTAAGGCAAGT 101755

QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
DB 101754 ATCGGTTATAAATCATCGCTTCATGTCGCGCAAGTGAAGAAATCGGTATGCTT 101695

QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuValLysLeuPheLysThrValLeu 100
DB 101694 ATTTCAACAGATCGTGGGATGTGGTGGTAAATGTTAATTTATTTCAAAACCACTT 101635

QY 101 AsnGluMetLysGluThrLysLysValSerValLysSerThrGlnLysLeuGlySer 120
DB 101634 AACCAATAAAATATGGAAGAACAAATATTTCTACAGATTTGGCTTAATAGGTCA 101575

QY 121 LysSerIleAsnPhePheGlnSerLeuGlyLysLeuValLysLeuPheLysThrValLeu 140
DB 101574 AAAGGATAGTATTTTCGTTCCCTTGGATTTAATATCAAAAGTCAGCTTCTGCTTTA 101515

QY 141 GlyAspThrProSerValGluThrLysLeuValLysSerValLysSerThrValLeu 160
DB 101514 GCGGATACGCGCTCTAGAACAAATTAATTTGGTGGCAATACAAATGTTGATGCTTAT 101455

QY 161 LysLysGlyGluValAspValLysLeuValLysThrGlnLysSerPheLysMetSer 180
DB 101454 CGTAATGGGAAATGATGCAAGTATATATGCAATCAATTAATTTGTTAATGATGCTG 101395

QY 181 GlnLysProValLysLeuLysLeuProLeuProGluLeuAspAsnGluLysGly 200
DB 101394 CAAAGCGTGTGTACAAATAGTCTTATCCAGAACTCAAGAACTCAAGAACTCAATTAAT 101335

QY 201 GluArgLysGlnValTrpAspTyrLysLeuProAspAlaLysValLysValLysAsn 220
DB 101334 GAAAGACACACAGACTGGGATATCTTATGAGCAGAACCAAAAGTACTATTAGATAGC 101275

QY 221 LeuLeuValArgTyrLeuGluSerGlnValLysLeuValLysLeuValLysLeuValLys 240
DB 101274 CTTTATGCTGTTATTTAGATGCTCAAAATTTATCAAGCGTGTGATGATGATGCTTCA 101215

QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaLysLeuValLysLeuValLysLeu 260
DB 101214 GAAACAGCGCTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 101155

QY 261 AsnGluLeuGlnLysValLysLeuValLysLeuValLysLeuValLysLeuValLys 280
DB 101154 AATGATCTGCGGTTGGTGTACAAAGCTGCTCAAGCAAGTATCAAAATGAATGAATGAAT 101095

QY 281 GluLeuValAlaGlyAlaAlaLysLeu 289
DB 101094 GAAATCGTAGCGGTGCGGCGGAT 101068

RESULT 7
AR274513_05/c
WFCOMMENT

Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name Begin End
AR274513_00 1 110000
AR274513_01 100001 210000
AR274513_02 200001 310000
AR274513_03 300001 410000
AR274513_04 400001 510000
AR274513_05 500001 610000
AR274513_06 600001 710000
AR274513_07 700001 810000
AR274513_08 800001 910000
AR274513_09 900001 1010000
AR274513_10 1000001 1110000
AR274513_11 1100001 1210000
AR274513_12 1200001 1310000
AR274513_13 1300001 1410000
AR274513_14 1400001 1510000
AR274513_15 1500001 1610000
AR274513_16 1600001 1710000
AR274513_17 1700001 1810000
AR274513_18 1800001 1910000
AR274513_19 1900001 2010000

Continuation (6 of 19) of AR274513 from base 500001 (AR274513 Sequence 1 from patent US

Alignment Scores:
Pred. No.: 1,42e-86 Length: 110000
Score: 1173.00 Matches: 221
Percent Similarity: 91.00% Conservative: 42
Best Local Similarity: 76.47% Mismatches: 26
Query Match: 82.09% Indels: 0
DB: 6 Gaps: 0

US-09-545-199F-4 (1-289) x AR274513_05 (1-110000)

QY 1 MetAlaGlyAlaLysGluThrLysLeuValSerValLysSerThrGlnLysLeu 20
DB 1934 ATGCGAGTGCCTGCAAGAGATACCAAAATTCGAGTACAGATACCAAAATC 1875

QY 21 ThrLysAlaMetGluMetValAlaLysSerValLysSerThrGlnLysMetSer 40
DB 1874 ACTAAGCAATGCAATGCGCAACCTGCAAAATCGTAAACGCAAGATCGTATGCT 1815

QY 41 SerSerArgProTyrSerGluThrLysLeuValLysSerValLysSerThrValLeu 60
DB 1814 GCATCTCGTCGATTCGAACTATCGTACGTTATGATGATGCTTAAGGCAAGT 1755

QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
DB 1754 ATCGGTTATAAATCATCGCTTCATGTCGCGCAAGTGAAGAAATCGGTATGCTT 1695

QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuValLysLeuPheLysThrValLeu 100
DB 1694 ATTTCAACAGATCGTGGGATGTGGTGGTAAATGTTAATTTATTTCAAAACCACTT 1635

QY 101 AsnGluMetLysGluThrLysLysValSerValLysSerThrGlnLysLeuGlySer 120
DB 1634 AACCAATAAAATATGGAAGAACAAATATTTCTACAGATTTGGCTTAATAGGTCA 1575

QY 121 LysSerIleAsnPhePheGlnSerLeuGlyLysLeuValLysLeuPheLysThrValLeu 140
DB 1574 AAAGGATAGTATTTTCGTTCCCTTGGATTTAATATCAAAAGTCAGCTTCTGCTTTA 1515

QY 141 GlyAspThrProSerValGluThrLysLeuValLysSerValLysSerThrValLeu 160
DB 1514 GCGGATACGCGCTCTAGAACAAATTAATTTGATGATGATGATGATGATGATGATGAT 1455

QY 161 LysLysGlyGluValAspValLysLeuValLysLeuValLysLeuValLysLeuValLys 180
DB 1454 CGTAATGGGAAATGATGCAAGTATATGCAATCAATTAATTTGTTAATGATGCTG 1395

QY 181 GlnLysProValLysLeuLysLeuProLeuProGluLeuAspAsnGluLysGly 200
DB 1394 CAAAGCGTGTGTACAAAGCTGCTCAAGCAAGTATCAAAATGAATGAATGAATGAAT 1335

QY 201 GluArgLysGlnValTrpAspTyrLysLeuProAspAlaLysValLysValLysAsn 220

```

Db      1334  GAAAGCAACAGACTGGGATATCTTTATGATCCAGAACCAAAAAGTACTATTAGATAGC 1275
Qy      221  LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluAsnLeuAlaSer 240
Db      1274  CTTTATGTTCTGTTATTAGAGTCCCAATTTATCAAGCGGTTGATAGATGATGAGTTCA 1215
Qy      241  GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
Db      1214  GAACAGCGGCTCGAATGGTAGCAATGAAGCAGCAACGAGTAATGCGAGTAAATTTAAT 1155
Qy      261  AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db      1154  AATGATCTGCGTGTGTTGATACACCAAGCTGTTCAAGCAAGTATCACAATGAATGAAT 1095
Qy      281  GluIleValAlaGlyAlaAlaIle 289
Db      1094  GAAATCGTAGCGGTGCGCGCGGAT 1068

RESULT 8
LOCUS      BD252190                      867 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Anti-bacterial vaccine compositions.
ACCESSION  BD252190
VERSION    BD252190.1  GI:33061960
KEYWORDS  JP 2002541790-A/74.
SOURCE    Actinobacillus pleuropneumoniae
ORGANISM  Actinobacillus pleuropneumoniae
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Actinobacillus.
1 (bases 1 to 867)
Lowery,D.E., Fuller,T.E. and Kennedy,M.J.
Anti-bacterial vaccine compositions
Patent: JP 2002541790-A/74 10-DEC-2002;
PHARMACIA AND UPJOHN CO
OS Actinobacillus pleuropneumoniae
PN JP 2002541790-A/74
PD 10-DEC-2002
PF 06-APR-2000 JP 2000611649
PR 09-APR-1999 US 60/128689,10-SEP-1999 US 60/153453 PI
DAVID E LOWERY,TROY E FULLER,MICHAEL J KENNEDY PC
C12N15/09,A61K39/02,A61K39/102,A61K39/395,A61P31/04,C07K14/285, PC
C07K16/12
PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/01,C12P21/02 PC
,C12P21/08,C12Q1/18,
PC GOIN33/15,GOIN33/50,GOIN33/53,GOIN33/569,GOIN33/577,C12N15/00,
PC C12N15/00,
PC C12N5/00
CC atpg
FH Key
FT CDS Location/Qualifiers
LOCUS      1. .867
/organism="Actinobacillus pleuropneumoniae"
/mol_type="genomic DNA"
/db_xref="taxon:715"

ORIGIN
Alignment Scores:
Pred. No.: 8.68e-86 Length: 867
Score: 1132.50 Matches: 215
Percent Similarity: 89.97% Conservative: 45
Best Local Similarity: 74.39% Mismatches: 28
Query Match: 79.25% Indels: 1
DB: 6 Gaps: 1

US-09-545-199F-4 (1-289) x BD252190 (1-867)

Qy      1  MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
Db      1  ATGGCAGGTGCGAAGAGATAGAACCAAAATTCGAAGTGTGAATAATCTCAAAAATC 60
Qy      21  ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40

```

```

Db      61  ACCAAAGCAATGGAAATGGTTGTACCTCTAAATATCGTTAAATCGTAAACGCAAGAGCGTATGGCT 120
Qy      41  SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db      121  GCCAGTCGCTCTTATTCGGAACCAATCCGTAAGTGATAGCCATATTTGCGAAGGAAGC 180
Qy      61  IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
Db      181  ATTGGTTATATAGCACCGGTTTTTAACGTGAACGCTGATATTAATAAAGATAGGCTATCTTGTTC 240
Qy      81  ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db      241  GTTTCGACCGATCGCGGTTTATCGCGTGGCTTTAATATCAATTTATTCAAGCGATTTG 300
Qy      101  AsnGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
Db      301  AATGAATTTAAACGCTGGAAGATAAAGACGTTAGTGTGAGCTTGCTTTAGTAGGTCG 360
Qy      121  LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db      361  AAAGCGGTAGCTTTTACCAAAATCTAGGCTTAACGTGAGATCTCAAGTAACGGGATTA 420
Qy      141  GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db      421  GCGGATAATCCGGAATGGAACGATATCGTGGCGCAGTTAATGAATGATTAATCGGTTTC 480
Qy      161  LysLysGlyGlnValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db      481  CGAACGGAGAGTGGATCGGTTTACGTCGGTTACACCGGTTTCAACCGTTTTGAAAAATACGATGCA 540
Qy      181  GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db      541  CAAAAACCTGTTATCGCACAGTTACTTCGTTACCTAACTAGATGACGATGAATTA--- 597
Qy      201  GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db      598  GATACGAAAGGTTCTAGGATATATTTATGAACCGAATCCACAAAGTTTATTGGATAGT 657
Qy      221  LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluAsnLeuAlaSer 240
Db      658  TTACTTGTTCGTTATTAGAACTCAGGTATACAGCGGTTGATAGATACTAGCTTCT 717
Qy      241  GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
Db      718  GAACAGCGGCTCGAATGGTAGCGATGAAGCGCAACAGATAATGCGGGTACATTATTC 777
Qy      261  AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db      778  GATGAATTACAAATTAGTGTATACAAAGCTCCGACAGCAATT 837
Qy      281  GluIleValAlaGlyAlaAlaIle 289
Db      838  GAAATGTTGCGGGTGGCGCAGCAATT 864

RESULT 9
AX553798
LOCUS      AX553798                      867 bp      DNA      linear      PAT 27-NOV-2002
DEFINITION Sequence 132 from Patent WO02075507.
ACCESSION  AX553798
VERSION    AX553798.1  GI:25897796
KEYWORDS
SOURCE    Actinobacillus pleuropneumoniae
ORGANISM  Actinobacillus pleuropneumoniae
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Actinobacillus.
1
Lowery,D.E., Fuller,T.E. and Kennedy,M.J.
Anti-bacterial vaccine compositions
Patent: WO 02075507-A 132 26-SEP-2002;
Pharmacia & Upjohn Company (US)
Location/Qualifiers
1. .867

```


Alignment Scores:

Pred. No.: 1.02e-81 Length: 304558
 Score: 1121.50 Matches: 212
 Percent Similarity: 89.97% Conservative: 48
 Best Local Similarity: 73.36% Mismatches: 28
 Query Match: 78.48% Indels: 1
 DB: 1 Gaps: 1

US-09-545-199F-4 (1-289) x AE017151 (1-304558)

QY	1	MetAlaGlyAlaLysGluLeuArgThrLysLeuAlaSerValLysSerThrGlnLysile	20
DB	7441	ATGCGAGGTGCTAAGAGATTAGAACCAAAATTCGAATTCGTAATACGCAAAATTT	7500
QY	21	ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer	40
DB	7501	ACGAACCAATGGAATGGTGCTACCTCTAAATTCGTAAACCGTAACGCAAGCGTATGCA	7560
QY	41	SerSerArgProTyrSerGluThrLysLeuArgAsnValLysSerHisValSerLysAlaThr	60
DB	7561	GCAGGTGCTCATATTACAGACGATCCGTAAAGTAATTAGCCATATCGCAAAAGGAAGC	7620
QY	61	IleGlyTyrLysHisProPheLeuValAspArgGluValLysValGlyMetIleVal	80
DB	7621	ATTGGTTATAGCACCCGCTTTTATAGCGGTGATGTTAAAGATTAAGGCTATTTAGTT	7680
QY	81	ValSerThrAspArgGlyLeuCysGlyLysLeuAsnValAsnLeuPheLysThrValLeu	100
DB	7681	ATCTCTACCGACCGCGTTTATGCGGTCTTATATATCAATTTATTCAGACACCTTG	7740
QY	101	AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer	120
DB	7741	AATGAATTTAAGCGTGGAAAGATAAGACGTCAGTGTGAATTTGGCTTAGTGGCTCA	7800
QY	121	LysSerIleAsnPhePheGlnSerLeuGlyLysLeuThrGlnAspSerGlyIle	140
DB	7801	AAGGAGTAGTCTTTTACCAATCTATGCTTAAAGTTAGGCTCATATTACAGGTTTA	7860
QY	141	GlyAspThrProSerValGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr	160
DB	7861	GGTATTCACCTGAAATGAAAGCTATTGTTGGTCAGTTAATGAATGAATTAATGCTTAT	7920
QY	161	LysLysGlyGluValAspValValTyrLysLeuValTyrAsnLysPheIleAsnThrMetSer	180
DB	7921	CGCAATGTTGAAGTGATATGTTTGTATGCTTATATCGTTTGAATATACAAATGCTCA	7980
QY	181	GlnLysProValLeuGlnLysLeuLeuProGluLeuAspAsnAspGluLeuGly	200
DB	7981	CAAAACACGTTATTGCAATATCTACCGTTACTTAAATTAGAAATGATGAATTA--	8037
QY	201	GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuAspAsn	220
DB	8038	GAACAAATGTTCTTGGATTATCTCTACGAGCTTAACCCACAGATTTATTGGATAGT	8097
QY	221	LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer	240
DB	8098	TTACTATTTCGTTATTAGAACAACAGGTATATCAAGCAGTTGTCGATAACCTTCTTCT	8157
QY	241	GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle	260
DB	8158	GAACAACCGCTCGAATGTTAGCAATGAGCGCCACACATATTCGGGAGCATTAATT	8217
QY	261	AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn	280
DB	8218	GATGATTACAAATTAGTGTATPACAAAGCTCGTCAAGCAAGTATTACAAATGAATTAAT	8277
QY	281	GluIleValAlaGlyValAlaAlaIle	289
DB	8278	GAAATTTGTCAGGTGCTGCTGCAATT	8304

RESULT 11

AX553832

LOCUS

866 bp DNA linear PAT 27-NOV-2002

DEFINITION

Sequence 166 from Patent WO02075507.

ACCESSION

AX553832 GI:25897830

VERSION

AX553832.1 GI:25897830

KEYWORDS

Mannheimia haemolytica

SOURCE

Mannheimia haemolytica

ORGANISM

Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 Pasteurellaceae; Mannheimia.

REFERENCE

1

Lowery, D.E., Fuller, T.E. and Kennedy, M.J.

Anti-bacterial vaccine compositions

Patent: WO 02075507-A 166 26-SEP-2002;

Pharmacia & Upjohn Company (US)

FEATURES

source

1..866
 /organism="Mannheimia haemolytica"
 /mol_type="unassigned DNA"
 /db_xref="taxon:75985"

gene

1..864

CDS

1..864

/gene="atpG"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD58199.1"
 /db_xref="GI:25897831"
 /translation="MAGKEIRTKIASVRNTQKITKAMEMVAASKRKTKOERMAASRP
 YAESRKAISHIAKNIIEYKHPFLTPRPVKVGLVSTDRGLCGGLNINLFTVLHE
 LKEKDDQVKGRLAVGVNGKISFNPWGLEIKGHINGLGTDPAMEDLVGVNGVNRVY
 REGEDIVVYVNRFINFMOKPTVQQLLPALENDSLEQTGSWDVLYEPNFCALHD
 SILVRYLESQVYQVDNLASEQARVAVKATNDAGNLINELQLVINKARQASITN
 ELNEIVAGAAAI"

ORIGIN

Alignment Scores:

Pred. No.: 3.44e-81 Length: 866
 Score: 1077.50 Matches: 208
 Percent Similarity: 87.89% Conservative: 46
 Best Local Similarity: 71.97% Mismatches: 34
 Query Match: 75.40% Indels: 1
 DB: 6 Gaps: 1

US-09-545-199F-4 (1-289) x AX553832 (1-866)

QY	1	MetAlaGlyAlaLysGluLeuArgThrLysLeuAlaSerValLysSerThrGlnLysile	20
DB	1	ATGCGAGGTGCTAAGAGATAAGAACCAAAATTCGAATTCGTAATACCAAAATTT	60
QY	21	ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer	40
DB	61	ACCAACGCGATGGAATGTTGCGCATCAAAATTCGTAATAACCCACGAGCGTATGCG	120
QY	41	SerSerArgProTyrSerGluThrLysLeuArgAsnValLysSerHisValSerLysAlaThr	60
DB	121	GCTTCTCGCCTTATGCTGAAAGTATTCGACGCAATTAGCCATATTGCCAAAGTAA	180
QY	61	IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal	80
DB	181	ATTGAGTATAACACCCATTTTGACCCACGTCGCGTAAACAAAAGTTGGCTATTAGTA	240
QY	81	ValSerThrAspArgGlyLeuCysGlyLysLeuAsnValAsnLeuPheLysThrValLeu	100
DB	241	GTTCACCATTCGCGTTTATGTGGTGGCTTAATCAATTTATTTAAACCGTTTAA	300
QY	101	AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer	120
DB	301	CATGAATGGAAGAAAGATGACCAAGGTGTTAAGTCTCGACITGCTGTGTCGGAAT	360
QY	121	LysSerIleAsnPhePheGlnSerLeuGlyLysLeuThrGlnAspSerGlyIle	140
DB	361	AAAGGATTCCTTTTAAACCAATGGGCGCTAGAGATTAAAGGTTCATATCAATGATG	420
QY	141	GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr	160

JOURNAL

Submitted (23-APR-2003) L. Prangeul, Institut Pasteur, Genopole, 25
rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail:
lfrangeul@pasteur.fr, lfrangeul@pasteur.fr

FEATURES

source

1. 349483

/organism="Photobacterium luminescens subsp. laumondii T101"
/mol_type="genomic DNA"

/strain="T101"

/db_xref="taxon:243265"

223. 1622

/gene="dnaA"

/locus_tag="plu0001"

223. 228

/gene="dnaA"

/locus_tag="plu0001"

234. 1622

/gene="dnaA"

/locus_tag="plu0001"

/codon_start=1

/transl_table=11

/product="Chromosomal replication initiator protein DnaA"

/protein_id="CAE12296.1"

/db_xref="GI:36783456"

/translation="MSLSLWQCLARLQDELPAEFPMWRPLQALSGNTLALYAPN
RFLVWVRKYNINGLNGLFCGAEVPLLRFEVGNKPVQNDSPQVQVTHPTPAPA
PQTSVRPMDTAVQPELSYNSNPNKHTDFNVEGSKQLARAAQVADNPAGAY
NPLFLYGGTGLKTHLHAGNSIMARKANAKVVMHSERFQDMVKALQVADNIAEFK
RYRSVDALLIDIQFFANKERSQBEFPHTEALLEGNOQIILTSRYPKEINGVEDR
LAKSRFGWLTVAIEPELETAVAILMKKADENEIQLPGEVAFI AKELRSNRELEGA
LNRVIANFTGRAITIDPVEALRDLALQELVTDNKTVAETIKIKVADLLSK
RRSRVARPRQAMALAKELTNHSLPEIGDAFGGRDHTTVLHACRKIEQREESHDIK
EDFSNLITLS"

1627. 2727

/gene="dnaA"

/locus_tag="plu0002"

1627. 2727

/gene="dnaA"

/locus_tag="plu0002"

/codon_start=1

/transl_table=11

/product="DNA polymerase III, beta chain"

/protein_id="CAE12297.1"

/db_xref="GI:36783457"

/translation="MKFTIEREQLLPLQVSSPLSGRPTLPILGNLLQVTEGSLLL
TGTDLEMMARVTLSPHEVGAATVPARFPIWGLPDGAEISVELDGRLLVRS
RSRSLTLPASDPNLDWQSEVERSPQATLRLKLIESTQFSAHQDVRVYLNGLF
ETEGEELTVAIDHRLAVCSMGIGQLPSHSVIVPRKGVIELMLLDGGDTPLQIQI
GNNIRAVGDFIFTSKLVGRFPDYRVLPRNPDKTLEANCMLKQAFSRAAILNE
AFRGVLYFSENQLRIITANNPEQEAEEIIVDSVYQGAENEIGFNVSVYLDVLTLCR
EVRLLTDAISSVQIEDCANHNAAYVVMPL"

2749. 3840

/gene="recP"

/locus_tag="plu0003"

2749. 3840

/gene="recP"

/locus_tag="plu0003"

/codon_start=1

/transl_table=11

/product="DNA replication and repair protein"

/protein_id="CAE12298.1"

/db_xref="GI:36783458"

/translation="MTLRLILDFRNIAADPLATGFNVLVSGSGKTSVLEAIV
TLGGRFRIOAGRVILHGCDEFVLHGRGQOENERSIGLSKNRNGDSKVIDGS
DGGKIALKMLPMQITPBGFTLNGGPKYRAFIDMGCFNRPFFSAMVNLKRL
KQNAALQVTRYSQLRPMQDELIPLANQINQWGEYVNTIQDITNCKQLPEFTL
SFSFOQGWKDSYALRQERDSTLTVASGPHKDLRIASGTVEMDLRSGQL
KLKMLARLQAGGYFRSQOQLYLLDDFASLDTSRQLLAKSTQAFVFSAI
NPQITMDLGNKSMFRVENGKIEVQFQD"

3860. 6326

/gene="gyrB"

/locus_tag="plu0004"

3860. 6274

/gene="gyrB"

/locus_tag="plu0004"

/codon_start=1

/transl_table=11

/product="DNA gyrase subunit B"

/protein_id="CAE12299.1"

/db_xref="GI:36783459"

/translation="MSNTYDSSSIKVLKGLDAVPRKPMYIGDTDGDTGLHVMFEV
DNAIDALAGHCSEVITIHADNSVSDQDGRGPIGTIHEEGVSAEAVIMTVLHAG
KEDNSYKVSGLHGVGVVVALSEKELVIRRDGVKHEQVTLGVQPLKVVGET
EQGTGRVFPWFMDFRNTDFQHDILAKRLSFLNSGVSIRLIDKRTNIEDHPHY
EGGKAFVEFLANKRTPIHPNVFYFTEKDGIGVEISQWNGDQFQENIYCFTNIPOR
DGTHLVGFRTAMRTILNSYMDKGVSKVSATGDDAREGLVAVISVVPKPKSS
QTKDLVSEVKTAVETLMNEKLYELLENDPAKTVVSKI IDAAAREARARSWT
RKGAIDLALGPKLADQERDPALELYLVEGDSAGSKQGNENKQAILPKSKI
LNVEKARDFKMLSSQEVATILTALCGIGEDVNPDKLRVHSIIMTADVDGSHKT
LLLTFFYRQPEIIIEHGHVFIAPPLLYKVKGQEQYIKODEANDYFQSLADGASL
VTNEHVPALGEPIELEKLVTFENGQYKIKRERLYPLSLNSLIYHNTEETLSDKA
KYENIEMLVTRLNDSQNSYSYLRHNERQMYEPVLCVTHGVDTYDYLDFDV
HGSEYHRTLLDGLKRLNIEDAYIERGERRPVSPFEQALNLTKESSRGLSIORYK
GLGENPQLWETTPETRRMVRVTVKDAIATDQLTTLMGDAVPRRAFTEENALK
AANTDI"

6300. 6326

/gene="gyrB"

/locus_tag="plu0004"

6415. 6926

/locus_tag="plu0005"

6415. 6894

/locus_tag="plu0005"

/function="Unknown"

/note="unnamed protein product; Similar to Hcp protein"

/codon_start=1

/transl_table=11

/protein_id="CAE12300.1"

/db_xref="GI:36783460"

/translation="MADMYMTIKKKQGLISAGCSSVDSIGNKYCANHFDOILVYSL
SHAITRTQNVHQPIIIQKPIDKSPLLGVASENELSCSIDFRTNSNGAQERYN
IKITGATITRISVLHPHNLNLPQESISLKYNITWNHIIAGTGGYSIWDDRIY"

6908. 6926

/locus_tag="plu0005"

/complement(8013..8778)

/locus_tag="plu0006"

/complement(8013..8768)

/locus_tag="plu0006"

/function="Unknown"

/note="unnamed protein product; Similar to unknown protein"

/codon_start=1

/transl_table=11

/protein_id="CAE12301.1"

/db_xref="GI:36783461"

/translation="MKQNDLKLKVELRELAKTSLFRYLTHSKIRMDPNSKINLVE
DILQVRIHCLSGNGAIEIQEIDHLKEQAFCLTANVKQYAIERKEKSYTNLI
LKOIGFVGQGTQIILAGYTVCKASGLACASGAPLMANGYNNVNGVYLLYRENG
GUREGYYVANKIGLSDDKADITYATVDVLVSIGYIFRKKVLKPREKSWLFRNINSDF
IRGWRMGVSLTTEGIVDVSVTLISYOLEKEK"

complement(8773..8778)

/locus_tag="plu0006"

/complement(8935..10056)

/gene="asd"

/locus_tag="plu0007"

complement(8935..10044)

/gene="asd"

/locus_tag="plu0007"

/transl_table=11

/product="aspartate-semialdehyde dehydrogenase (ASA)

dehydrogenase) (ASADH)"

/protein_id="CAE12302.1"

/db_xref="GI:36783462"

/translation="NMKNVGFIRGWMGVSLMQRMIEQDFVICPFFFTSHQGP
APDFTGQQTQLQNAFDIHALGQDIIISCGQDYTNIEIYPLKRLTQGWYIDAAAL
RNDSDIIILDPVNHTHIQOGLNKIKITFVGNCITSLMLSLGGLFANDLVMSVA
TYQASAGAGARHRELLVQMSLHTQVAKELQDPASAILDIERKVDTFTRSGVMSTEK

```

FSVPLAGSLIPWIDKQNSREBWKGAETNKILNTGNNIITVDGLCYRIGALRCH
SOAFLTKKDISIPEIROLIAAHNDWVRVLPNDRELSMRELTFAAVTGILDPVGL
RLNMGPEVLSAFTVGDQLLWGAEPILRMLRIIL"
RBS
complement(10051. .10056)
/gene="asd"
/locus_tag="plu0007"
complement(10251. .10472)
/gene="ogrK"
/locus_tag="plu0008"
complement(10251. .10472)
/gene="ogrK"
/locus_tag="plu0008"
/locus_tag="plu0008"
/function="Unknown, probable positive regulator of phage
P2 late gene transcription"
/codon_start=1
/transl_table=11
/product="Ogr protein"
/protein_id="CAE12303.1"
/db_xref="GI:36783463"
/translation="MIKPLCGKAHARSFEHSCHTKERYNQCNINGCATFVSHT
FVTRISRPGEVISVKPHPKTKTQLNVT"
complement(10549. .11634)
/locus_tag="plu0009"
complement(10549. .11634)
/locus_tag="plu0009"
/function="Unknown, probable tail synthesis"
/note="unamed protein product; Similar to protein D (gpd)
of Enterobacteria phage P2"
/codon_start=1
/transl_table=11
/protein_id="CAE12304.1"
/db_xref="GI:36783464"

Alignment Scores:
Pred. No.: 8.07e-72 Length: 349483
Score: 1004.00 Matches: 195
Percent Similarity: 84.08% Conservative: 48
Best Local Similarity: 67.47% Mismatches: 44
Query Match: 70.26% Indels: 2
DB: 1 Gaps: 1

US-09-545-199F-4 (1-289) x BX571859 (1-349483)

QY 1 MetAlaGlyAlaLysGluLeuThrLysIleAlaSerValLysSerThrGlnLysIle 20
Db 37140 ATGCGCGCGCCAAAGAGATACGTACCAAGATCGCCAGGTGCTCAAAACACCAAAAATC 37081
QY 21 ThrLysAlaMetGluMetValAlaLysSerLysMetArgLysThrGlnGluArgMetSer 40
Db 37080 ACTAAGACGATGAGATGGTGTGCGTCCAAATGCGTAAACGCGATCGCATGCGG 37021
QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db 37020 GCCAGCGCGCTTATGACAAACCATACGACGCGATGTCACCTTGCCTTAGGPAAT 36961
QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
Db 36960 CTGGAAATAAACAACCGTACCTTGAAGAGCGTCAAAACCAACGTCGGTATCTGTT 36901
QY 81 ValSerThrAspArgGlyLeuCysGlyGlyValLeuAsnValAsnLeuPheLysThrValLeu 100
Db 36900 GTTTCACCGATCGTGGCTGTGCGTGGTGTGATCTACTATCTGTTTCAAAAACGCTT 36841
QY 101 ArgGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
Db 36840 TCAGAAATGAAGACTGTGTGATAAAGATCCAGTGTGAGTGGCGCTTATCGCTCT 36781
QY 121 LysSerIleAsnPheGlnSerLeuGlyLysIleLysIleLeuThrGlnAspSerGlyIle 140
Db 36780 AAGCTACTCTTCTTCTGCTTCTGTTGGGGGATACGTTTGTCTCAGGTAAACGCGATG 36721
QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160

```

```

Db 36720 GGAGATAACCCCTTCACTGCTGATTTGATCGGGCCAGTGATATCATCTGTCGGGGATAC 36661
QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 36660 GATGAGGACGTCGTGATAAACTGTATGTGTGACAAACAAGTTCATCAATACAAATGTC 36601
QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db 36600 CAGGAACCGACTATTACTCAGTTTATACCTTCTGCTCCCGAGATGATGAACACTG--- 36544
QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db 36543 ---AAGAAAGAAATCCTGGGATATTATACGAACCTGATCCTTAAGCGTGTCTAGATATA 36487
QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValAlaValGluAsnLeuAlaSer 240
Db 36486 CTGCTGCGTGTATGTAGATCGAGTTATCAGGGCGTGTGAAACCTGGGTAGT 36427
QY 241 GluGlnAlaLysArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
Db 36426 GAACAGCCCGCAAGATGTTGGCGATGAAAGCGCAACTGATATGTTGGTGGACCTGATC 36367
QY 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db 36366 AAGAGTTGCGATGTTGTTTATACAAAGCTCGTCAGGCCAGCATAACTCAGGAGCTGACC 36307
QY 281 GluLeuValAlaGlyAlaAlaIle 289
Db 36306 GAGATTGTCTCGGTGCTTCCGCGTT 36280
RESULT 14
AE004342/c
LOCUS
AE004342 12928 bp DNA linear BCT 10-OCT-2003
DEFINITION
Vibrio cholerae O1 biovar eltor str. N16961 chromosome I, section
250 of 251 of the complete chromosome.
ACCESSION
AE004342 AE003852
VERSION
AE004342.1 GI:9657358
KEYWORDS
SOURCE
Vibrio cholerae O1 biovar eltor str. N16961
ORGANISM
Vibrio cholerae O1 biovar eltor str. N16961
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
REFERENCE
1 (bases 1 to 12928)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Uitterback, T., Fleischmann, R.D.,
Nierman, W.C. and White, O.
DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae
JOURNAL
Nature 406 (6795), 477-483 (2000)
MEDLINE
20406833
PUBMED
10952301
REFERENCE
2 (bases 1 to 12928)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Uitterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
Direct Submission
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1. .12928
/organism="Vibrio cholerae O1 biovar eltor str. N16961"
/mol_type="genomic DNA"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:243277"
/chromosome="I"

```

```

/notes="biotype: E1 Tor"
/136..2307
/gene="VC2758"
CDS
/136..2307
/gene="VC2758"
/notes="similar to GB:M87049 SP:P1177 GB:M59368 GB:M74164
PID:145900; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="fatty oxidation complex, alpha subunit"
/protein_id="AAF95897.1"
/db_xref="GI:9657358"
/transl_table=11
/translation="MIYQATLQVQLANGIAELSCAPASVKNLKLHLSLKDALD
ALAADSVKGLSSDEKAEIVGADITEFLGFAKEAELEDEWLFANRINFKLELDL
FPTLSUKHGLGGCECVIATDFRIGDATTIGLPTKLGIMPGFGGTVALPRLIGA
DSAMEITQKACRAEALAVGLLDAIVDSKLIDSALTTLTQAEELDKWKRQOK
TSALTLSKLEAMNFTMAKMGVAQVAKHPAPMTSVVTIEAASLPDAALDIERKH
FIKLAKSTEALQVGLPLNDQYIKGLAKOSAKAASQDQTHRAVLGAGIMGGGIAYQSA
LKGVPVLMKDAIAPHSLKELVTEAAKILLKQLERKIDGFKMAGILASITPSLHVAGID
QADVIVBAVVENPKVAVLSEVEGLVDTELTISNTIPINLLAKSLKRPQFQGM
HFNPNHMEPLVLEIIGETISEDTINRVVAYAAKMGKSPVVDNCPGFVNRVLPVPF
AGSLLMRUGANFTEIDKMERQFGMPMPGAYLLDVGVIGIDTAHQAQVMAEGFFTRA
KSGREIDALYEAKFGKNGSGFYQYTVDKKGPKAQFSDVLAILAPVCGAQPSDF
QOULIERTMIPMEVYVLCLEGIITASAQEDMALVYGLGPPRPGGVFRVLDITGIA
NYVMAEKYADLALYQVQLLKNVAQQTSTFYSAAQQYSAL"
/2333..3496
/gene="VC2759"
CDS
/2333..3496
/gene="VC2759"
/notes="similar to GB:M87049 SP:P21151 GB:M59368 GB:M74164
PID:145901; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="fatty oxidation complex, beta subunit"
/protein_id="AAF95898.1"
/db_xref="GI:9657360"
/transl_table=11
/translation="MNTVVIVDCLTPMRSKGGAFRQRAEDLSAHLMKGILARNPQ
VNPKEIDYWCQVQTLQEQFNVARNAALAGLPIBIGAVTVNRLOGSMOALHNDAA
RAIMVGDABEILCVGVEHMGVPMTHVDHPGSKNVAKAAGMGLTASMLKGLHGI
SRQODEFAARSHARAAHATEGRPKNELITEGHRAADGTLFTLIDIVEIRPITVAG
LAELRPVDPANGTITACTSNLSDGASAMLVMSQKAKALGLTIRARIKAMAVAGD
PSIMGYGVPVATKALQAGITMQDMVVELNEFAAQSILPCAKDLGLEMMDDKYNL
NGGALAHGPGCSGSTRITSLINLMEAKDARYLATMCIQLGQGIATIPER"
/complement(3563..4471)
/gene="VC2760"
CDS
/complement(3563..4471)
/gene="VC2760"
/notes="similar to SP:P37598 GB:U00096 PID:1742732
PID:1742742 PID:178949; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, LysR family"
/protein_id="AAF95899.1"
/db_xref="GI:9657361"
/transl_table=11
/translation="MMPKSKSLEMLDAVARLGSTFAAEVLEHKVPASISYVRQTESE
LGVLPRLPKVELITPAGELFMLEAQLQLRQEEIRATQKAAHGWKRLKVLTDNV
VGLDKMPMEAFQITDFEALQINMEVNGSWEAIAQGRADIVIGATAVFPVGDFFE
VRDGLIDWAFVMSPNPCVREQNLSGFTISQFLAICLDDTSSVLPKHTHYPKORR
LLKPNWYSALSKSLGGLGVMPRHMAQPLLAGSQVLEKVPDELKLSHCLLWREKOD
NKKLIMWVNYLGSQDLYSDMLQAK"
/4600..5829
/gene="VC2761"
CDS
/4600..5829
/gene="VC2761"
/notes="similar to GB:U00096 PID:1742733 PID:1742743
PID:178950; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="multidrug resistance protein"
/protein_id="AAF95900.1"
/db_xref="GI:9657362"
/transl_table=11
/translation="MNFDFGHEKMKVSPFQLVYLSALSMGLFIATDMYLPFAKGMETA
FATGPEQIALSLTVFLVGNALGQLLWGLSKDFEHRNTLIAGLLVFTLASAGLPSST
WQOLLRLRPTQAIQVCAPIVQCAVIVQWVSGQQLFATIMPLVLSALPOLGV
LLADYFQWHSIFLTLTLMGIVLVSTGQTNEIAEPTTS: AQDILKLLKTYLGNV
MMFACASARFAITLTGMPETMAQGLGYQDGLSIFIPOTIAPVGVGLKKAQVQYGD
ELVKQILALFTLSAAMI: FVASQWTLSSIPFLAPFCLIAVANGAMYIIVNRLSSA
KQSPATAAGLQNSLQSSLSALVAANASQAQPVGTGIAVLCTGGLWGLVLSNRT
LREQFTTPNRSRVSDSE"
/complement(5891..7252)
/gene="VC2762"
CDS
/complement(5891..7252)
/gene="VC2762"
/notes="similar to SP:P17114 PID:290578 PID:43267
PID:290579 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylglucosamine pyrophosphorylase"
/protein_id="AAF95901.1"
/db_xref="GI:9657363"
/transl_table=11
/translation="MKFSVTILAAKQTRMHSNMPKVLHTLAGKPMVKHVIDTCNNLG
ADNIELVTHGGDQMCQALANENVNWLQQLGTGHAVDQASPHFODDEKILVLYGD
VELISEDITELLEAQTGDIAGLLTVLEDTGYGRIVRKGQVPAIVEOKDASERQK
LIKEYNTGVVATGTRDLKRWLAGLNNNAQGEVYLTQVIAAHDGEGAVZAVPHSHSI
EYGVNDRIQARLERAFQARQAKLLEQGMLEDPARFDRGLTQCGSDVEIDVNVV
IEGVNIGNNVIGAGSILKDCEDDNTVIRPVSIEGATVGENCTVPTFRLRPGAE
LRDDAHVGNFVEMKNALGEGSKANHLTYLGDAIGKGVNAGVITCNVDGANKHT
VIGDDVFGSDCQLVAPVTIGNGATIGAGTTLTKNVAEGLVITRAPERKIAGWQRPAA
KKK"
/complement(7499..7921)
/gene="VC2763"
CDS
/complement(7499..7921)
/gene="VC2763"
/notes="similar to SP:P2988 PID:48340; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, epsilon subunit"
/protein_id="AAF95902.1"
/db_xref="GI:9657364"
/transl_table=11
/translation="MAATFHLVDVSAEKKLSGLVETPQVTGSEGEIGIPHGHTPLL
TAKPGMVRIVKLHGHEEYVSGGIVEVQPGTATVLADTAIRGEDLDAKAEAKRR
ABERILNQHDMQFAQAASELAKAIAQLRVIELTKRR"
/complement(7940..9343)
/gene="VC2764"
CDS
/complement(7940..9343)
/gene="VC2764"
/notes="similar to GB:X16050 SP:P12986 PID:48339;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, beta subunit"
/protein_id="AAF95903.1"
/db_xref="GI:9657365"
/transl_table=11
/translation="WATGKIVQIICAVVDVBPQSEVSVVDALNVDSKERLVLEVO
QOLGGVIRAIWNGSSDGLRGMVQNTGAPISVPGTKTLGRIMNVLGDAIDERGDI
GAEVYSIHRPAPSIEBSSATLETTGVKVIDLICFPKAGKIGLFGGAGVKTVMV
MELINNALVALGSLGSVFAVGERTREGDNFHEMQEAGVNVNVEQPSALGRMVSQVNV
EPFGNRLVALGTLMAEKFRDEGRDVLFF:DNIRYVTLAGTEVSALLGRMVSQVNV
PILAEIMGVQBERITSTKKGSITTSQAVVYVADDTDPSPATTFADLDAIVLNRNLA
AMGLYPAIDPLDST:RQLDPLVWQGEHYDVARGVQATLQVRKEIKDLIALGMDLSE
ADQVVARARKIERELTQPVHVAEVTGDPGVYVFLKETLAGFKGLLAGDYDILPEQA
FNYVCGTIDDAENAKKL"
/complement(9378..10244)
/gene="VC2765"
CDS
/complement(9378..10244)
/gene="VC2765"
/notes="similar to SP:P12990 PID:48338; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, gamma subunit"
/protein_id="AAF95904.1"
/db_xref="GI:9657366"

```

```
/translation="MAGAKBIRTKIGSVKSTOKITKAMENVAASVRRSQDAMESRRP
YAOTIRKIVGHVANASLEVRHPVLEERBAKRVYIIISDRGLCGGLNINLPKKAITD
MQWKEKGAQIELAIIQSKATAFNNSGAKVAQVSGDGLSDLEDLIGSGVGLMKRY
DGLGRLILVFNQFVNTWOKPKIIDLPLPKSDSDENQDRHMDYIYEPEPKLLD
ALLURFLESQVYQGVENLACEQAARVAMKAATDNASNLIDLQLVYNKARQAATQ
ELSEIVGGAAGV"
gene      complement(10289..11860)
CDS       /gene="VC2766"
          complement(10289..11860)
          /gene="VC2766"
          /note="similar to SP:PI2985 P1D:48337; identified by
          sequence similarity; putative"
          /codon_start=1
```

Alignment Scores:

Pred. No.:	3.94e-73	Length:	13928
Score:	998.50	Matches:	188
Percent Similarity:	83.74%	Conservative:	54
Best Local Similarity:	65.05%	Mismatches:	46
Query Match:	69.87%	Indels:	1
DB:	1	Gaps:	1

US-09-545-199F-4 (1-289) x AE004342 (1-12928)

QY	1	MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
DB	10244	ATGCCCGCGCAAGAGATACGTACTAGATCGGAGGTGTGAAAGCACTCAGAAGATT 10185
QY	21	ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
DB	10184	ACGAAAGCGATGAATGTGGCTGTTCAAAAATGCGTCTCAGACGAGCTATGGAG 10125
QY	41	SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
DB	10124	TCATCAGCGCCTTACCACAAACATACATAAGTATGATCGGTGATCGCCACGCAAGT 10065
QY	61	IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
DB	10064	CTAGATGATCGTCATCGTATCTAGAGAGCGTGAAGCAGCGGATGGTTTACATCAT 10005
QY	81	ValSerThrAspArgGlyLeuGlyLysGlyLysLeuAsnValAsnLysThrValLeu 100
DB	10004	ATTCTACAGACCGCGTTTGTGTGTGTGTGTGAACATTAATCTGTTCAAAAAGCCATC 9945
QY	101	AsnGluMetLysGluTrpLysGlyLysAspValSerValGlnLysSerLeuIleGlySer 120
DB	9944	ACGGATATGCAACTTGGAAAGAGAAAGTGGCGAGATTGAGCTGGCGGATATCGGCTCG 9885
QY	121	LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
DB	9884	AAAGCGACCGCTTTTTCACACACAGCGCGGAAAGTACGGGCTCAGGTTTCTGGTTG 9825
QY	141	GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
DB	9824	GCGCACGCCCAAGCTTAGAGATCTGATCGGTTGCGTTCATGCGTGAAGAATAAC 9765
QY	161	LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
DB	9764	GATAAAGGTGATGATAGACCGCTCTATCTGTGTTCACACAGTTTGAACACCATGGTG 9705
QY	181	GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAsnAspGluLeuGly 200
DB	9704	CAAAACCAAGATCGATCAATGCTACCTTTGCCAAATCGGACACGCAAGATATG--- 9648
QY	201	GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
DB	9647	CAGCGTGATCATGTGGGACTACATCTACGAGCGTGAACCAAGGCTCTATTAGATGCG 9588
QY	221	LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
DB	9587	CTTTTACTGCGCTTCATGTGATCGAGGTTTATCAAGGTTTATCAAAAACCTAGCGTGT 9528
QY	241	GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeulle 260

Db	9527	GAGCAACGACCCGAAATGGTTGGATGAAGCTGTACCGATACCGGAGCAACCTGATT 9468
QY	261	AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
DB	9467	GATGATTTACAATCTGTGTACACAAAGCCGTCAGGCGCGGATTTACACAGAATTGTCG 9408
QY	281	GluIleValAlaGlyAlaAlaAlaIle 289
DB	9407	GAATCGTAGGCGCGCGCTGCGGTT 9381

RESULT 15

BX950851.50/c

WPCOMMENT

Sequence split into 51 fragments LOCUS BX950851 Accession BX950851

Fragment Name	Begin	End
BX950851_00	1	110000
BX950851_01	100001	210000
BX950851_02	200001	310000
BX950851_03	300001	410000
BX950851_04	400001	510000
BX950851_05	500001	610000
BX950851_06	600001	710000
BX950851_07	700001	810000
BX950851_08	800001	910000
BX950851_09	900001	1010000
BX950851_10	1000001	1110000
BX950851_11	1100001	1210000
BX950851_12	1200001	1310000
BX950851_13	1300001	1410000
BX950851_14	1400001	1510000
BX950851_15	1500001	1610000
BX950851_16	1600001	1710000
BX950851_17	1700001	1810000
BX950851_18	1800001	1910000
BX950851_19	1900001	2010000
BX950851_20	2000001	2110000
BX950851_21	2100001	2210000
BX950851_22	2200001	2310000
BX950851_23	2300001	2410000
BX950851_24	2400001	2510000
BX950851_25	2500001	2610000
BX950851_26	2600001	2710000
BX950851_27	2700001	2810000
BX950851_28	2800001	2910000
BX950851_29	2900001	3010000
BX950851_30	3000001	3110000
BX950851_31	3100001	3210000
BX950851_32	3200001	3310000
BX950851_33	3300001	3410000
BX950851_34	3400001	3510000
BX950851_35	3500001	3610000
BX950851_36	3600001	3710000
BX950851_37	3700001	3810000
BX950851_38	3800001	3910000
BX950851_39	3900001	4010000
BX950851_40	4000001	4110000
BX950851_41	4100001	4210000
BX950851_42	4200001	4310000
BX950851_43	4300001	4410000
BX950851_44	4400001	4510000
BX950851_45	4500001	4610000
BX950851_46	4600001	4710000
BX950851_47	4700001	4810000
BX950851_48	4800001	4910000
BX950851_49	4900001	5010000
BX950851_50	5000001	5064019

Continuation (51 of 51) of BX950851 from base 5000001 (BX950851 Erwinia carotovora subsp.

Alignment Scores:

Pred. No.:	8.22e-72	Length:	64019
Score:	993.00	Matches:	194
Percent Similarity:	82.70%	Conservative:	45

Best Local Similarity: 67.13% Mismatches: 48
Query Match: 69.49% Indels: 2
DB: 1 Gaps: 1

US-09-545-199F-4 (1-289) x BX950851_50 (1-64019)

QY 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
DB 56259 ATGCCCGCGCCAAAGAGATACGTAGTAAGTCGCAAGCGTCCAAATACGCAAGATC 56200

QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
DB 56199 ACCAAAGCAATGGAATGTCGCGCTCTAAATGCTAAATCGCAGGAGGTATGCGG 56140

QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
DB 56139 GCCAGCGCTCTTATGCGGAACACATACCAATGTGTCACCTTCGCTAGGGAAT 56080

QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysValGlyMetIleVal 80
DB 56079 CTGGAATATATAACACCGTACCTGGAAGAACGTGGCGTTAAGCGGTTGGGTATTGGTG 56020

QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
DB 56019 GTGTCTACTGACCGTGGCTTGTGTGTTGTAACATTAACCTGTTCAAAAACTGCTG 55960

QY 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
DB 55959 GCTGATATGAATCCTGGAGCGCAAGGCGTTGAACGTGATTAGCGCTGATTGGTTCC 55900

QY 121 LysSerIleAsnPhePheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
DB 55899 AAAGCGCTTCTTCTTCGGTTCGGTAGCGGAACATTTGTCTCAGGTTACCGGTATG 55840

QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
DB 55839 GGGGATAACCTTCGGTATCAGAAATTGTCGCCCGGTTAAAGTTATGCTGCAAGCCTAT 55780

QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
DB 55779 GACGAGGTCTGTCGCAAGCTGATATATGTAAGCAACAAAGTTTATCAATACCATGCT 55720

QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
DB 55719 CAGGAACCGCTTATTGTTCAAGTGTTACCGTTACCGCTTCAGATGACAGTGAGTTG--- 55663

QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
DB 55662 ---AAGAAAGAAAGCCCTGGGATTACTGTATGACCCGACTCTAAGTCGCTGGATACC 55606

QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
DB 55605 CTGCTGCCCGTTATGTAGATCTCAGGTTTATCAGGGCGCTGTAGAAATCTGGCTAGT 55546

QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
DB 55545 GAGCAGCGCGCGCAATGGTTGGCATGAAAGCGCGACTGATAACGGCGTAGCCTGATC 55486

QY 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGluAlaSerIleThrAsnGluLeuAsn 280
DB 55485 AAAGAGCTTCAGTTGGTATACACACAGCGCGCTCAGCCAGCACTTACTCAGGAACCTACC 55426

QY 281 GluIleValAlaGlyAlaAlaIle 289
DB 55425 GAAATCGTCGGGGAGCGCTCCGGGTT 55399

Search completed: November 9, 2004, 14:48:48
Job time : 24192 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 9, 2004, 02:54:00 ; Search time 629 seconds
(without alignments)
2411.896 Million cell updates/sec

Title: US-09-545-199F-4
Perfect score: 1429
Sequence: 1 MAGAKEIRTKIASVKSTQKI.....ARQASITNELNIVAGAAAI 289

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Zgapop 6.0 , Zgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/US070.spool_p/US09545199/runat_08112004_112526_20941/app_query.fasta_1.455
-DB=N_Geneseq_238sep04 -Qfmt=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOPCPL=0
-LOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09545199 @CGN 1.1.885 @runat_08112004_112526_20941 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_238sep04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	1429	100.0	1972 3	AAC79582 Virulence
2	1429	100.0	1972 6	ABQ83459 Pasteurel
C 3	1173	82.1	110000 2	Continuation (5 of
C 4	1173	82.1	110000 2	Continuation (6 of
5	1132.5	79.3	867 3	AAC79654 Virulence
6	1132.5	79.3	867 6	ABQ83531 Actinobac

ALIGNMENTS

RESULT 1
AAC79582
ID AAC79582 standard; DNA; 1972 BP.
XX
AC AAC79582;
XX
DT 08-FEB-2001 (first entry)
XX
DE Virulence gene #2.
XX
KW Virulence gene; antibacterial; vaccine; bacterial infection; septicemia;
XX
OS Pasteurella multocida.
XX
PN WO200061724-A2.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009218.
XX
PR 09-APR-1999; 99US-0128689P.
XX
PD 10-SEP-1999; 99US-0153453P.
XX
(PHAA) PHARMACIA & UPJOHN INC.
XX
Lowery DE, Fuller TE, Kennedy MJ;
XX
WPI: 2000-647422/62.
XX
P-FSDB; AAB44522.

7	1077.5	75.4	866	6	ABQ83549
8	1004	70.3	864	10	ACF70377
9	1004	70.3	876	10	ADf02992
10	1004	70.3	110000	10	ACF65385_5
11	1004	70.3	110000	10	ACF67367_34
12	991	69.3	864	5	AAH81345
13	978	68.4	894	11	ACH97650
14	926.5	64.8	876	11	ABD03905
15	926.5	64.8	1050	11	ABD03801
16	894	62.6	879	9	ADA29481
17	885	61.9	930	12	ADL03242
18	885	61.9	94750	4	AAF28551
19	804	56.3	110000	6	ABA32787_0
20	760	53.2	873	10	ABZ38350
21	753	52.7	46593	3	AAA81456
22	753	52.7	110000	3	AAA81489_5
C 23	753	52.7	349980	3	AAF21612
C 24	750	52.5	3125	5	AA878013
25	750	52.5	3125	5	AA886688
26	739	51.7	3936	5	AA882564
27	702	49.1	2786	5	AA893030
28	673	47.1	3091	5	AA892566
29	650.5	45.5	531	3	AA879635
30	650.5	45.5	531	6	ABO83512
31	616	43.1	2583	5	AA889854
32	530.5	37.1	915	10	ADH84612
33	530.5	37.1	7750	2	AAH13153
34	530.5	37.1	7750	6	ABS98948
35	524.5	36.7	4815	2	AAV29571
36	523.5	36.6	110000	6	ABA90521_18
37	518.5	36.3	879	6	ABN68209_08
38	518.5	36.3	110000	6	ABN71527_08
39	518	36.2	851	6	ABX72809
40	510.5	35.7	903	10	ADC94003
41	508.5	35.6	873	6	ABN68210
42	504.5	35.3	2207	2	AAV29572
43	501.5	35.1	876	10	ABX07252
44	501.5	35.1	888	8	ABZ42460
45	501.5	35.1	10399	2	AAV52293

Abq83549 Pasteurel
Acf70377 Photorhab
Adf02992 Bacterial
Continuation (6 of
Continuation (35 o
Aah81345 Escherich
Ach97650 Klebsiell
Abd03905 Pseudomon
Abd03801 Pseudomon
Ada29481 DNA encod
Adl03242 DNA encod
Aaf28551 Genomic f
Aba32787 Buchnera
Abz38350 N. gonorr
Aas81456 N. mening
Continuation (6 of
Aaf21612 Neisseria
Aas78013 DNA encod
Aas86688 DNA encod
Aas82564 DNA encod
Aas93030 DNA encod
Aas82566 DNA encod
Aac79635 Virulence
Abq83512 Actinobac
Aas89854 DNA encod
Adh84612 Enterococ
Aax13153 Enterococ
Abs98948 Enterococ
Aav29571 L. lactis
Continuation (19 o
Abn68209 Streptoco
Continuation (9 of
Abk72809 Bacillus
Adc94003 E. faeciu
Abn68210 Streptoco
Aav29572 L. lactis
Abx07252 S. pneumo
Abz42460 Streptoco
Aav52293 Streptoco

XX Attenuated Pasteurellaceae bacteria comprising mutations in virulence
PT genes, useful as a live attenuated vaccine against bacterial infections.
XX
FS Claim 1; Page 67-68; 322pp; English.
XX
CC The family Pasteurellaceae encompasses several pathogens that infect a
CC wide variety of animals. The present invention relates to virulence genes
CC from Pasteurellaceae. The present sequence is one such virulence gene.
CC The present sequence may be mutated in order to produce an inactive gene.
CC The inactive virulence gene may in turn be used to produce a vaccine,
CC which is useful for treating bacterial infections such as septicemias,
CC bronchopneumonias, rhinitis and wound infections
XX
SQ Sequence 1972 BP; 652 A; 312 C; 438 G; 570 T; 0 U; 0 Other;

Alignment Scores:		
Pred. No.:		1972
Scores:	1.88e-144	289
Percent Similarity:	129.00	
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	100.00%	Mismatches: 0
DB:	3	Indels: 0
		Gaps: 0
US-09-545-199F-4 (1-289)	x AAC79582 (1-1972)	

QY	1	MetAlaGlyValAlaLysGluLeuIleArgThrIysLysIleAlaSerValLysSerThrGlnLysIle	20
Db	364	ATGGCAGGTGCTAAAGAGATAGAACCAAAATCGGAGTGTAAAGAGTACACAAAATT	423
QY	21	ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer	40
Db	424	ACTAAGCGATGGAAATGGTTGCTGCCTCGAAAATGGTAAACGCCAAGAACGCATGTCT	483
QY	41	SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr	60
Db	484	TCTTCACGCCCTATTTCAGAACCAATACGTAACTGATAGCCACAGCTTCCAAAGCAACG	543
QY	61	IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal	80
Db	544	ATTGTTTCAACAGCATCCATTTTATAGTGAATCGCGAAGTAAAAAAGTGGGCATGATGTT	603
QY	81	ValSerThrAspArgGlyLeuCysGlyLeuAsnValAsnLeuPheLysThrValLeu	100
Db	604	GTGTCCACAGATCGTGTCTTTGTGTGGCTTAAACGTGAACCTGTGTTTAAACTGTTATTA	663
QY	101	AsnGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer	120
Db	664	AATGAAATGTAAGAAATGGAAGAAAAAGATGTTTCGGTTCAATGTAGTTTAAATCGTTCT	723
QY	121	LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle	140
Db	724	AAATCTATCAACTTTTCCAACTTTGGGAATTAATAATTTAACCCAGATTCAGGTATT	783
QY	141	GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr	160
Db	784	GGTGATATCCCTCTGTTGACGAGTTAATGTGTTCACTCATATCTATGATTCATCTAT	843
QY	161	LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer	180
Db	844	AAAAAGGGGAAGTAGATGTGTGTATTAGTTTATAACAAATTTATTAAACAGATGTGC	903
QY	181	GlnLysProValLeuGlnLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly	200
Db	904	CAAAAGCCAGTATTCGAAAAATTAATTCATTACCAGAAATAGATATAATGATGAATTAGGC	963
QY	201	GluArgLysGlnValTyrTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn	220
Db	964	GAAAGAAAAACAAGTTGGGATTATTTACGAACCTGATCGAAGATTTATTAGATTAAT	1023
QY	221	LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer	240
Db	1024	TTATTGGTTGTTATTAGATCTCAGGTTTATCAGCGAGTTGAAACTTGGCTTCT	1083

Qy	241	GlutInAlaAlaAtgMetValAlaMetLysAlaAlaThrAspAsnAlaClYAsnLeuIle	260
Db	1084	GAGCAAGCGGCTCGANTGGTCGCCATGAAGCAGCAACGATACCGAGTAACTTAAT	1143
Qy	261	AsnGluLeuGlnLeuValTyrAsnLysAlaAtgGlnAlaSerIleThrAsnGluLeuAsn	280
Db	1144	AATGAGTTACAGTTAGTCTATACAAAGCTCGTCAGCAAGTATTACAAATGAATTAAT	1203
Qy	281	GluIleValAlaGlyAlaAlaAlaIle	289
Db	1204	GAATGTGTCCGGTCAGCAGCAATT	1230

RESULT 2	
ABQ83459	
ID	ABQ83459 standard; DNA; 1972 BP.
XX	
XX	
AC	ABQ83459;
XX	
XX	
DT	07-AUG-2003 (revised)
DT	24-JAN-2003 (first entry)
XX	

XX	Pasteurella multocida atpG gene SEQ ID NO:3.
DE	Antibacterial; vaccine; gram negative bacterial virulence gene;
XX	identification; virulence; Pasteurellaceae; gene; ds.
KW	Pasteurella multocida.
OS	WO200275507-A2.
XX	26-SEP-2002.
XX	17-JAN-2002; 2002WO-US001971.
XX	15-MAR-2001; 2001US-00809665.
XX	(PHAA) PHARMACIA & UPJOHN CO.
XX	Lowery DE, Fuller TE, Kennedy MJ;
PI	WPI; 2002-740868/80.
XX	P-PSDB; ABP54474.
XX	New mutant gram-negative bacteria, useful as vaccines and for identifying
XX	new anti-bacterial agents that target virulence genes and their products.
XX	Claim 1; Page 72-73; 350pp; English.

The present invention describes a gram-negative bacteria comprising a mutation in a gene, where the mutation results in decreased activity of a gene product encoded by the mutated gene. Also described is a method for producing a gram-negative bacteria mutant or an attenuated Pasteurellaceae bacteria. The mutated genes have antibacterial activity and can be used in vaccines. The gram-negative bacteria or the attenuated Pasteurellaceae bacteria can be used as vaccines in the fields of human medicine or veterinary medicine, and for identifying new antibacterial agents that target the virulence genes and their products. ABQ83458 to ABQ83578 and ABP54473 to ABP54551 represents sequences used in the exemplification of the present invention. (Updated on 07-AUG-2003 to correct OS field)

Sequence 1972 BP; 652 A; 312 C; 438 G; 570 T; 0 U; 0 Other;

```

Alignment Scores:
  Seq No.:      Length:      1972
  Score: 1.88e-144
  Matches: 1429.00
  Mismatches: 289
  Conservative: 0
  Mismatches: 0
  Indels: 0
  Gaps: 0
  Gaps: 0

```

US-09-545-199F-4 (1-289) x ABQ83459 (1-1972)

QY 1 MetAlaGlyAlaLysGluLeuLeuArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
Db 364 ATGGCAGGTGCTAAAGAGATAAGAACCAAAATCGAGTGTAAAGATACACAAAAAT 423
QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
Db 424 ACTAAGCGATGGAATGGTGTGCTCCGAAATGCGTAAACGCGAAGACGCAATGTCT 483
QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db 484 TCTTCACCCCTTATTAGAAACAATACGTAACGTGATTAGCCACGTTTCCAAAGCAACG 543
QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
Db 544 ATGGTTACAGCATCCATTTTAGTGGATCGGCAAGTAAAGAAATGGCGCATATTGTT 603
QY 81 ValSerThrAspArgGlyLeuCysGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db 604 GTGTCCACAGATCGTGTCTTTGTGGTGGCTTAAACGTGAACCTGTTTAAACTGTATTA 663
QY 101 AsnGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
Db 664 AATGAATGAAGAAATGGAAGAAAGATGTTTCGTTCAATTGAGTTTAAATCGGTCT 723
QY 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db 724 AATCTATCAACTTTTCCAACTTTTGGCAATTAAATTTTAAACCAAGATTCAGGTATT 783
QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaThr 160
Db 784 GGTGATACCTCTCTGTTGAGCAGTTTAAATGTTTATATACAAATTTATACACGATGTCG 843
QY 161 LysGlyGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 844 AAAAAAGGGAAGTATGTTGTTTATTTAGTTTATATACAAATTTATACACGATGTCG 903
QY 181 GlnLysProValLeuGlnLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db 904 CAAAGCCAGTATTCGAAAAATTAATTCATACCAAGTATGATAATGATGATGATGATG 963
QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAsn 220
Db 964 GAAAGAAACAAAGTTGGGATTATTTACCAACTGATCGCAAGTATTTATGATAAT 1023
QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
Db 1024 TTATCGTTCGTTTATTTAGAAATCTCAGTTTATCAAGCAGCTTGAAAACCTTGCTTCT 1083
QY 241 GluGlnAlaAlaArgMetValAlaMetValAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
Db 1084 GAGCAAGCCGCTCGAATGGTCGCATGAAAGCAACACAGATAACCGATTAATTTAAT 1143
QY 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db 1144 AATCAGTTACAGTTAGTCTATATAACAAAGCTCGTCAAGCAAGTATTACAAATGAATTAAT 1203
QY 281 GluIleValAlaGlyAlaAlaIle 289
Db 1204 GAAATTCGTGCGGTGCGAGCGCAATT 1230
RESULT 3
AAT42063_04/c
Continuation (5 of 19) of AAT42063 from base 400001 (Haemophilus influenzae complete genome)
Fragment Name Begin End
AAT42063_01 100001 110000
AAT42063_02 200001 310000
AAT42063_03 300001 410000
AAT42063_04 400001 510000
AAT42063_05 500001 610000
AAT42063_06 600001 710000

WP AAT42063_07 700001 810000
WP AAT42063_08 800001 910000
WP AAT42063_09 900001 1010000
WP AAT42063_10 1000001 1110000
WP AAT42063_11 1100001 1210000
WP AAT42063_12 1200001 1310000
WP AAT42063_13 1300001 1410000
WP AAT42063_14 1400001 1510000
WP AAT42063_15 1500001 1610000
WP AAT42063_16 1600001 1710000
WP AAT42063_17 1700001 1810000
WP AAT42063_18 1800001 1930121
Alignment Scores:
Pred. No.: 2,31e-114 Length: 110000
Score: 1173.00 Matches: 221
Percent Similarity: 91.00% Conservative: 42
Best Local Similarity: 76.47% Mismatches: 26
Query Match: 82.09% Indels: 0
DB: 2 Gaps: 0
US-09-545-199f-4 (1-289) x AAT42063_04 (1-110000)
QY 1 MetAlaGlyAlaLysGluLeuLeuArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
Db 101934 ATGGCAGGTGCTAAAGAGATAAGAACCAAAATCGAGTGTACAAAGATACACAAAAATC 101875
QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
Db 101874 ACTAAGCGAATGGAATGGTGGCAACCTCGAAATGCGTAAACGCGAGATCGTATCGCT 101815
QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db 101814 GCATCTCGTGTATCTGAACTATTCGTAACTTATTAGTCTCATGTGTAAAGCAAGT 101755
QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
Db 101754 ATCGTTATAACATCCGTTCTTAGTTAGCGGGAAGTGAAGAAATCGGTATCTTGGTT 101695
QY 81 ValSerThrAspArgGlyLeuCysGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db 101694 ATTTCAACAGATCGTGGGATGCTGGTGGTTAAATGTTAAATTTATTTCAAAACCACTT 101635
QY 101 AsnGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
Db 101634 ACCAAATAAAATTCGAAAGAACAAATATTTCTACAGATTGGGCTTAATAGTTCTA 101575
QY 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db 101574 AAAGGATTAATTTTTCGTTCTTGGATTATATATCAAAAGGTCAAGCTTTCTGTTTA 101515
QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaThr 160
Db 101514 GCGATACGCGCTCTAGAAATTAATTTGGTGGCAATACAAATGTTGATCTTAT 101455
QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 101454 CGTAATGTGAATGTACAGTTTATATTCATACAAATAAATTTGTTAAATACGATGTCG 101395
QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db 101394 CAAAAGCTGTTGTACAACTAGTTCTTTACAGAAATCTAAAGACGATCATTTAAAT 101335
QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db 101334 GAAAGCAACAGCTTGGGATATCTTTATGAGCGAGAACCAAAAGTACTATTAGATAGC 101275
QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluAsnLeuAlaSer 240
Db 101274 CTTTATGTTCTGTTATTTAGATGCTCCAAATTTATCAAGCGTTGTAGATATAGTTCTA 101215
QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260

Db 101214 GAACACGCGGCTCGAATGGTAGCAATCAAGACGACGACGATAATCCAGGTAATTAATT 101155

Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280

Db 101154 AATGATCTGGGTGGGTGTACACAAAGCTCGTCAAGCAAGTATCACAATGAATGAAT 101095

Qy 281 GluIleValAlaGlyAlaAlaAlaIle 289

Db 101094 GAAATCGTAGCGGTGGGGCGGCGATT 101068

RESULT 4

AAT42063_05/c

Conjunction (6 of 19) of AAT42063 from base 500001 (Haemophilus influenzae complete genome sequence split into 19 fragments LOCUS AAT42063 Accession AAT42063

WP	Fragment Name	Begin	End
WP	AAT42063_00	1	110000
WP	AAT42063_01	100001	210000
WP	AAT42063_02	200001	310000
WP	AAT42063_03	300001	410000
WP	AAT42063_04	400001	510000
WP	AAT42063_05	500001	610000
WP	AAT42063_06	600001	710000
WP	AAT42063_07	700001	810000
WP	AAT42063_08	800001	910000
WP	AAT42063_09	900001	1010000
WP	AAT42063_10	1000001	1100000
WP	AAT42063_11	1100001	1200000
WP	AAT42063_12	1200001	1300000
WP	AAT42063_13	1300001	1400000
WP	AAT42063_14	1400001	1500000
WP	AAT42063_15	1500001	1600000
WP	AAT42063_16	1600001	1700000
WP	AAT42063_17	1700001	1800000
WP	AAT42063_18	1800001	1830121

Alignment Scores:

Pred. No.: 2,31e-114 Length: 110000

Score: 1173.00 Matches: 221

Percent Similarity: 91.00% Conservative: 42

Best Local Similarity: 76.47% Mismatches: 26

Query Match: 82.09% Indels: 0

DB: 2 Gaps: 0

US-09-545-199F-4 (1-289) x AAT42063_05 (1-110000)

Qy 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20

Db 1934 ATGCGAGTSCAAAGAGATAAACCAAAATTCGACAGTGTACAAAGTACACAAATAATC 1875

Qy 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnLysArgMetSer 40

Db 1874 ACTAAGCAATGGAATGGTGGCACTCGAATAATCGGTAAACCCAGAGTCGTATGGCT 1815

Qy 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60

Db 1814 GCATCTGTCGCGTATCTGAAACTATCCGTAACGTTATAGTCATGTCATAGGCAAGT 1755

Qy 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysValGlyMetIleVal 80

Db 1754 ATCGGTATAAACATCCGTTCTTAGTTGAGCGCAAGTGAAGAAATTCGGTATCTTGGT 1695

Qy 81 ValSerThrAspArgLysGlyLeuGlyGlyLysValAsnValAsnLeuPheLysThrValLeu 100

Db 1694 ATTTCAACAGATCGGGGATGTGGGTGTTAAATGTTAATTTATTTCAAAACCACTT 1635

Qy 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120

Db 1634 AACCAAAATATAAAATTTGGAAGACAAATATTTCTACAGATTTGGGCTTTAATAGTTCA 1575

Qy 121 LysSerIleAsnPhaPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140

Db 1574 AAAGGATAGTTTTCCTGCTTCTTGGATTAAATCAAGGTCAGCTTCCTGTTTGA 1515

Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160

Db 1514 GGGATACCCCGCTCTAGAGATTAATTCGTGGCAATACAAATCTTTGATGCTTAT 1455

Qy 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180

Db 1454 CGTAATGGTGAATTCAGTTCAGTTTATTCATACAAATAAATTTGTTAATACGATGCG 1395

Qy 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnGluLeuGly 200

Db 1394 CAAAAGCCCTGTTGTACAAATAGTTCCTTTACCAGAAATCTAAAGACGATCATTTAAAT 1335

Qy 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220

Db 1334 GAAAGACACAGACTTGGGATTCCTTTATGAGCAGACCAAAAGTACTATTAGATAGC 1275

Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluLeuLeuAlaSer 240

Db 1274 CTTTATGTCGTTATTTAGAGTCCCAATTTATCAAGCGTTGTAGATATGTTAGTTCA 1215

Qy 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260

Db 1214 GAACACGCGCTCGAATGGTAGCAATGAAAGCAGCAACGATAATGCAGTAATTTAAT 1155

Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280

Db 1154 AATGATCTGGGTGGTGTACAAAGCTCGTCAAGCAAGTATCACAATGAATTTGAAT 1095

Qy 281 GluIleValAlaGlyAlaAlaAlaIle 289

Db 1094 GAATCTGACCGGTGGCGGCGGATT 1068

RESULT 5

AAC79654

ID AAC79654 standard; DNA; 867 BP.

XX AC AAC79654;

XX DT 08-FEB-2001 (first entry)

XX DE Virulence gene #61.

XX XW Virulence gene; antibacterial; vaccine; bacterial infection; septicemia; bronchopneumonia; rhinitis; wound infection; SS.

XX OS Actinobacillus pleuropneumoniae.

XX DN WO200061724-A2.

XX PD 19-OCT-2000.

XX PF 06-APR-2000; 2000WO-US009218.

XX PR 09-APR-1999; 99US-0128689P.

XX PR 10-SEP-1999; 99US-0153453P.

XX PA (PHAA) PHARMACIA & UPJOHN INC.

XX XX Lowery DE, Fuller TE, Kennedy MJ;

XX DR WPI; 2000-647422/62.

XX XX P-PSDB; AAB44579.

XX PT Attenuated Pasteurellaceae bacteria comprising mutations in virulence genes, useful as a live attenuated vaccine against bacterial infections.

XX PS Claim 7; Page 290-291; 322pp; English.

XX CC The family Pasteurellaceae encompasses several pathogens that infect a wide variety of animals. The present invention relates to virulence genes from Pasteurellaceae. The present sequence is one such virulence gene. The present sequence may be mutated in order to produce an inactive gene. The inactive virulence gene may in turn be used to produce a vaccine,

-----CCGAAAGGCTGATAGCCATATTCGGAAGGAAGC 180

Fri Nov 12 17:22:00 2004

us-09-545-199f-4.p2n.rng

QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
 Db 181 ATTGGTTATAGACCCCGTTTAACTGAACGATGATTAATAAAAGTACGCTATCTTGC 240
 QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
 Db 241 GTTTCGACCGATCGCGTTTATGCGGTGCTTAATATCAATTTATCAAGCGACTTG 300
 QY 101 AsnGluMetLysGluTyrLysGluLysAspValSerValGluLeuSerLeuIleGlySer 120
 Db 301 AATGAATTTAAACCTGGAAGATAAAGCGTTAGTGTGAGCTTGGTTAGTAGGTCG 360
 QY 121 LysSerIleAsnPheGlnSerLeuGlyLysIleLeuThrGlnAspSerGlyIle 140
 Db 361 AAAGCGTAAGCTTTTACCAAAATCTAGCTTAACCTGAGATCTCAAGTAACGGGATTA 420
 QY 141 GlyAspThrProSerValGluLeuLeuGlySerValAsnSerMetIleAspAlaTyr 160
 Db 421 GCGGATATCCGAAATCGAAGATGATCGTGGCGCAGTTAATGAATGATTAATGCGTTC 480
 QY 161 LysLysGlyGluValAspValValTyrLeuValTyrLeuLysPheIleAsnThrMetSer 180
 Db 481 CGAAACGGGAAGTGGATCGCGTTTACGCTTACACCGCTTTTGAATACGATGTC 540
 QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnGluLeuGly 200
 Db 541 CAAAACCTGTTATCGCACAGTTACTCGGTACCTAACTAGATGACGATGAATTA-- 597
 QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
 Db 598 GATACGAAAGGTTCTGCGGATTAATTTATGAAACCGAATCCACAACTTTTATGATAGT 657
 QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
 Db 658 TTACTTGTTCGTTATTAGAACTCAGGTATACCAAGCAGTTGTAGATAAAGCTAGCTCT 717
 QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
 Db 718 GAACAAAGCCGCTCGAATGTTAGCGATGAAGCGGCAACAGATTAATGCGGTCACATTAATC 777
 QY 261 AsnGluLeuGlnLeuValTyrLeuLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 Db 778 GATGAATTAATAGTTGATTAACAAGCTCGCAAGCAAGCATTAACAATGAATTAAC 837
 QY 281 GluLeuValAlaGlyAlaAlaIle 289
 Db 838 GAAATTTGTCGGGTGCGCGCAAT 864

RESULT 7
 ABQ83549
 ID ABQ83549 standard; DNA; 866 BP.
 XX AC ABQ83549;
 XX AC
 DT 29-AUG-2003 (revised)
 DT 24-JAN-2003 (first entry)
 DE Pasteurella haemolytica atpG gene SEQ ID NO:166.
 XX Antibacterial; vaccine; gram negative bacterial virulence gene;
 KW identification; virulence; Pasteurellaceae; gene; db.
 XX Mannheimia haemolytica.
 OS
 XX WO200275507-A2.
 PN
 XX PD 26-SEP-2002.
 XX PF 17-JAN-2002; 2002WO-US001971.
 XX PR 15-MAR-2001; 2001US-00809665.
 XX PHAA) PHARMACIA & UPJOHN CO.

XX Lowery DE, Fuller TE, Kennedy MJ;
 XX WPI; 2002-740868/80.
 XX P-PSDE; ABP54547.
 XX New mutant gram-negative bacteria, useful as vaccines and for identifying
 XX PT new anti-bacterial agents that target virulence genes and their products.
 XX Claim 5; Page 329-330; 350pp; English.
 XX The present invention describes a gram-negative bacteria comprising a
 XX mutation in a gene, where the mutation results in decreased activity of a
 XX gene product encoded by the mutated gene. Also described is a method for
 XX producing a gram-negative bacteria mutant or an attenuated
 XX Pasteurellaceae bacteria. The mutated genes have antibacterial activity
 XX and can be used in vaccines. The gram-negative bacteria or the attenuated
 XX Pasteurellaceae bacteria can be used as vaccines in the fields of human
 XX medicine or veterinary medicine, and for identifying new antibacterial
 XX agents that target the virulence genes and their products. ABQ83458 to
 XX ABQ83578 and ABP54473 to ABP54551 represents sequences used in the
 XX exemplification of the present invention. (Updated on 29-AUG-2003 to
 XX standardise OS field)
 XX SQ Sequence 866 BP; 283 A; 150 C; 195 G; 238 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6.92e-107 Length: 866
 Score: 1077.50 Matches: 208
 Percent Similarity: 87.89% Conservatives: 46
 Best Local Similarity: 71.97% Mismatches: 34
 Query Match: 75.40% Indels: 1
 DB: Gaps: 1
 US-09-545-199F-4 (1-289) x ABQ83549 (1-866)
 QY 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
 Db 1 ATGGCAGGTGCTAAGAGATAGAACCAAAATTCGTAAGTCTTCGTAATACACAAAAAT 60
 QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnLysArgMetSer 40
 Db 61 ACCAAGCGATGGAATGCTGTCGCGCATCAAAATGCTGTAACCAAGCGATGCGG 120
 QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
 Db 121 GCTTCTCGCCTTATGCTGAAAGTATTGCGAAGCAATTAGCCATATTGCCAAGGTAA 180
 QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
 Db 181 ATTGAGTATTAACACCCCAATTTTACCCCGCTCGGTAAAGAGTTGCTATTAGTA 240
 QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
 Db 241 GTTTCACCGCATCGCGTTTATGTCGCTTAATATCAATTTATTAACACCGTTTA 300
 QY 101 AsnGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
 Db 301 CATGATTAAGAAAGAAAGATGACCAAGGTGTTAAGTCTCGACTTCTGTCGCGGAAT 360
 QY 121 LysSerIleAsnPheGlnSerLeuGlyLysIleLeuThrGlnAspSerGlyIle 140
 Db 361 AAAGGATCTCTTTTAAACCAATGGGCTAGAGATTAAGATTAAGGTTCATATCGGATT 420
 QY 141 GlyAspThrProSerValGluLeuLeuGlySerValAsnSerMetIleAspAlaTyr 160
 Db 421 GGTGATACACCGCAATGGAAGATTAGTCGGTATTGTTAATGATGTTAATGATGCTAC 480
 QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
 Db 481 CGTGAAGCGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnGluLeuGly 200

Db 541 CAAAACCCGACAGTCAACAGTTCCTTCTTGGCTGCCTGCACTGGAAATGACTCATTA--- 597
 QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
 Db 598 GAGCAAACTGGTCTCTGGGATTCCTATGAACCAAAATCCACAAGCGGTATTAGACAGC 657
 QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaLaValGluAsnLeuAlaSer 240
 Db 658 TTTACTGGTTCGTATTAGATCTCAAGTTATCAGCGAGTGGTAGATAATCTTGGCTCT 717
 QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuLe 260
 Db 718 GAACAGGCTGCTCGAATGCTGCAATGAAGCAGCAACCGATAACGCGAGGTAAATCTGATT 777
 QY 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 Db 778 AATGAGTTACAGTTAGTGTATACCAAGCTCGTCAAGCAAGTATACGAATGAATTAAT 837
 QY 281 GluileValAlaGlyAlaAlaLaile 289
 Db 838 GAAATTTGCGGGTGGCGCAATTT 864

RESULT 8

ID ACF70377 standard; DNA; 864 BP.
 AC ACF70377;

20-NOV-2003 (first entry)

Photorhabdus luminescens nucleotide sequence #8844.

Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 detection; food; gene expression; plant; animal; microorganism; toxin;
 antibiotic; biopesticide; virulence factor; disease model; plague;
 whooping cough; gene; ds.

Photorhabdus luminescens.

WO200294867-A2.

28-NOV-2002.

07-FEB-2002; 2002WO-1E003040.

07-FEB-2001; 2001FR-00001659.

(INSP) INST PASTEUR.

(CNRS) CNRS CENT NAT RECH SCI.

Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 Buchrieser C;

WFI; 2003-148459/14.

Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 useful e.g. as therapeutic antimicrobials and agricultural pesticides.

Claim 2; SEQ ID NO 8844; 1205pp; French.

The invention relates to the isolation of genes and their encoded
 proteins from Photorhabdus luminescens. The isolated sequences are
 sources of probes and primers for detecting the genome of P. luminescens
 and related species; to study polymorphisms; for gene analysis and for
 detection/amplification of the genes. Antibodies (Ab) raised against the
 polypeptides encoded by the genes are used for detection/identification
 of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 carry a gene-containing vector are used to select compounds that
 modulate, regulate, induce or inhibit expression of the genes in plants,
 animals or microorganisms other than P. luminescens and are able to alter
 response or sensitivity to toxins and antibiotics produced by P.
 luminescens. Cells transformed to express the genes are useful for

CC recombinant production of the proteins, particularly toxins and
 antibacterials useful as insecticides, bactericides and fungicides. The
 genes, proteins, vectors containing the genes and Ab are also useful
 therapeutically (to treat microbial infection by bacteria or fungi that
 are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 biopesticides. Other uses of the genes and the proteins are as virulence
 factors and for identifying targets of human diseases for which P.
 luminescens is a model (particularly plague and whooping cough). This
 sequence represents one of the isolated P. luminescens genes

SQ Sequence 864 BP; 236 A; 183 C; 229 G; 216 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.25e-99 Length: 864
 Score: 1004.00 Matches: 195
 Percent Similarity: 100.0% Conservative: 48
 Best Local Similarity: 84.08% Mismatches: 44
 Query Match: 70.26% Indels: 2
 DB: 10 Gaps: 1

US-09-545-199F-4 (1-289) X ACF70377 (1-864)

QY 1 MetAlaGlyAlaLysGluileArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
 Db 1 ATGGCGCGCGCAAAAAGAGATACGTACCAAGATCGCCAGCGTGCACAAAACGCAAAAATC 60
 QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
 Db 61 ACTAAAGCGATGGAGATGGTTGCTGCGTCCAAAATGCGTAAACGCGAGGATCCATGGCG 120
 QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
 Db 121 GCCAGCGCTCCTTATGCGAGAAACCATACGCGAGCGGTATGCTCACCCTGGGTAGGTAAT 180
 QY 61 IleGlyTyrLysHisProPheLeuValAspArgLysValLysValGlyMetIleVal 80
 Db 181 CTGGAATAAACAATCCGTCACCTTGAAGAGCGTGAACCAACGTCGCGGTATCTCGGT 240
 QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
 Db 241 GTTTCTACCGATCGTGGCTTGTGGGGTGGTTGAATACTAATCTGTTCAAAAAATCGCT 300
 QY 101 AsnGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
 Db 301 TCAGAAATGAAGACTGGTCTGATAAAGATGTCAGTGTGAGTGGCGCTATTCGGCTCT 360
 QY 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
 Db 361 AAGGTACTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420
 QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
 Db 421 CGAGATAACCCCTTCACCTGCTGAATTCATCGCGGCGAGTGAATATCATCTGCGCGGCATC 480
 QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
 Db 481 GATGAGGAGCGTCTGGATAAACTGTATGTGTGTGACAAACAAGTTCATCAATACATGTC 540
 QY 181 GlnLysProValLeuGluLysIleProLeuProGluLeuAspAsnAspGluLeuGly 200
 Db 541 CAGGAACCGACTATTACTCAGTTATTAATCTCTGCTGCGGAGATGATGAACACATC- 597
 QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
 Db 598 ---AGAGAAATCTCTGGGATTAATTTATACCAACTGATCTAAGCGGTGCTAGATATA 654
 QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaLaValGluAsnLeuAlaSer 240
 Db 655 CTGCTGGTCTGTTATAGAAATCGCAGGTTTATCAGGCGCGTGGTTGAAAACCTGGCTAGT 714
 QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuLe 260
 Db 715 GAACAGGCGCGCAAGAAATGGTGGCGGATGAAGCCCAACTGATATATGGTGGCAGCTGATC 774

QY 261 AsnGluLeuValThrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 DB 775 AAGAGTTGCAATGGTTATACCAAGCTCGTCAGGCCACGATACTCAGAGCTGACC 834
 QY 281 GluLeuValAlaGlyAlaAlaLaille 289
 DB 835 GAGATTGCTCGGTGCTTCGCGGTT 861

RESULT 9

ADP0292
 ID ADP0292 standard; DNA; 876 BP.
 AC ADF0292;
 AC ADF0292;
 DT 12-FEB-2004 (first entry)
 XX Bacterial polynucleotide #3277.
 XX Proteus mirabilis infection; bacterial infection; antibacterial;
 KW immunostimulant; gene; ds.
 XX Proteus mirabilis.
 OS US6605709-B1.
 XX 12-AUG-2003.
 XX 05-APR-2000; 2000US-00543681.
 XX 09-APR-1999; 99US-0128706P.
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX Breton GL;
 XX WPI; 2003-895291/82.
 DR P-PSDB; ADF07164.

XX New Proteus mirabilis polypeptides and polynucleotides, useful as
 PT reagents for diagnosis of bacterial disease, as components of
 PT antibacterial vaccines, as targets for antibacterial drugs, or as
 PT biocontrol agents for plants.
 XX Disclosure; SEQ ID NO 3277; 870pp; English.
 XX The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunising an individual against *P. mirabilis*, a
 CC method for evaluating a compound for the ability to bind a *P. mirabilis*, a
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular
 CC targets for diagnosing, preventing and treating pathological conditions
 CC resulting from bacterial infection, as reagents for diagnosis of
 CC bacterial diseases, as components of antibacterial vaccines, as targets
 CC for antibacterial drugs or as bio-control agents for plants. This
 CC sequence represents a Proteus mirabilis polynucleotide of the invention.

XX SQ Sequence 876 BP; 261 A; 170 C; 229 G; 216 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6.37e-99 Length: 876
 Score: 1004.00 Matches: 194
 Percent Similarity: 84.08% Conservative: 49
 Best Local Similarity: 67.13% Mismatches: 44
 Query Match: 70.26% Indels: 2
 DB: 10 Gaps: 1

US-09-545-199f-4 (1-289) x ADF0292 (1-876)

QY 1 MetaAlaGlyAlaLeuGluLeuArgThrLysIleAlaSerValLysSerThrGlnLysIle 20

DB 13 ATGCCGCGCGCAAAAGAGATACGTTCCAGATGCCAGTGTGCCAAACACACAGAGATC 72
 QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
 DB 73 ACTAAGCGATGGAGATGGTCGCGCGTCCGAAATGCGTAAACACGAGAACCCATGGCA 132
 QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
 DB 133 GCCAGCGCTCCTTATGAGAAACCATCGCAGATGTTGTTGCTCACCTGCGTTAGGAAAT 192
 QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
 DB 193 CTGGATATAAATCATCTCAAGAGCGGTGAAGTTAAACGTGTGGGTACCTGGTT 252
 QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
 DB 253 GTTTCGACCGACCGTGGTTGTGTGGTGTGTAACATTAACTTGTTCAAAAAAGCTG 312
 QY 101 AsnGluMetLysGluTyrLysGluLysValSerValGlnLeuSerLeuIleGlySer 120
 DB 313 GCAGACATGAAGAGTGTCTGATAAAGGTGAAGATGATTTAGCCCTTGTGTGTTCT 372
 QY 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
 DB 373 AAAGCGTTTCATTCTTCTGCTCTGTAGTGGGAATGTTGTCGTCAGTAACAGGTATG 432
 QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
 DB 433 GGTGATGATCCGCAACTATCTGACCTTATCGGCCCGATCAATATCATGTGCGAGGCTTAT 492
 QY 161 LysLysGlyGluValAspValValLysLeuValLysValLysValLysValLysVal 180
 DB 493 GACGAAGTCTGTTAGATAAACTGTATGTGTGGTGGCAATAAAGTTTCATCAATACCATGGCC 552
 QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
 DB 553 CAAGAGCTTAAATTCGCAAGTTTACCAATACCAAGGTCATGATGATGAAGAAGTAA 612
 QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
 DB 613 GAAAAATCT-----TGGGATATCTGACGAACCTGATCCTTAAGACGTTACTGGATCT 666
 QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluLeuAlaSer 240
 DB 667 TTAAGTGTGCTGTTATATGTAATGTCAGGTTTATCAGAGCGTCTAGAGAACTTAGCGAGT 726
 QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
 DB 727 GAGCAGCAGCAGCAAGATGGTTCGGATGAAGCCGCTACAGATTAACGGTGTATCTGATT 786
 QY 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 DB 787 AAAGAACTGCAATTTGGTGTACAAACAAAGCAGCGCAAGCAAGCAATTAAGCAAGTACA 846
 QY 281 GluLeuValAlaGlyAlaAlaLaille 289
 DB 847 GAAATTTGATCTGTCGTCAGCGCGGTTA 873

RESULT 10

ACF65385_5
 Continuation (6 of 7) of ACF65385 from base 500001 (Photographus luminescens nucleotide
 WP Sequence split into 7 fragments LOCUS ACF65385 Accession ACF65385
 WP Fragment Name Begin End
 WP ACF65385_0 1 110000
 WP ACF65385_1 100001 210000
 WP ACF65385_2 200001 310000
 WP ACF65385_3 300001 410000
 WP ACF65385_4 400001 510000
 WP ACF65385_5 500001 610000
 WP ACF65385_6 600001 618776

Alignment Scores:
 Pred. No.: 4.56e-96 Length: 110000


```

Db 66149 ACTAAGCGATGAGATGGTGTGCTGCTCAAAATCGTAAACCGAGATCGCATGCGC 66208
Qy 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db 66209 GCCAGCGCTCTTATGACAGAACCATACGACGCGTGTGCTACCTGCGTGGTGGTAAAT 66268
Qy 61 IleGlyTyrIleHisProPheLeuValAspArgGluValLysValGlyMetIleVal 80
Db 66269 CTGGAATATAAACATCCGACCTTGAAGAGCGTGAACCAACCGTGTGCGGTATCGGTT 66328
Qy 81 ValSerThrAspArgGlyLeuCysGlyValLeuAsnValAsnLeuPheLysThrValLeu 100
Db 66329 GTTCTACCGATCGTGGCTGTGCGGTGTGTTGAATACCTATCTGTTCAAAAGCTGTT 66388
Qy 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
Db 66389 TCAGAATGAAGAACTGGTCTGATAAAGATGCCAGTGTGAGCTGGCGCTTATCGGCTCT 66448
Qy 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db 66449 AAGCTACCTCTTCTTCTGCTCTGTTGGGTTACGTTGTGCTCAGTACAGGCAATG 66508
Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaThr 160
Db 66509 GGAGATAACCTTCACTGCTGAATGTATGTCGGCCAGTGAATATCATGCTCGGCGCATC 66568
Qy 161 LysLeuGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 66569 GATGAGGACCTGTGATAACTGATGTGTTGTCGACAAACAGTTCAATACATGTCTCC 66628
Qy 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnGluLeuGly 200
Db 66629 CAGGAACCGCATATTACTCAGTTATTACCTCTGCTGCGGAGATGATGAACACTG-- 66685
Qy 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAsn 220
Db 66686 ---AAGAAGAAATCCCTGGATATTATTATACGAACTGATCTTAAGCGGTGCTAGATATA 66742
Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
Db 66743 CTGCTCGCTGCTTATGATAGATCGCAGGTTTATCAGGCGCTGTTGAACCTGGCTAGT 66802
Qy 241 GluGlnAlaAlaArgMetValAlaValLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
Db 66803 GAACAGCGCCGACGATGTGCGGATGAAGCCGCACTGATATGTTGGTGGCAGCTGATC 66862
Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db 66863 AAAGAGTTGCAGTTGTTTATACAAAGCTGTCAGGCCAGCATAACTCAGGAGCTGACC 66922
Qy 281 GluIleValAlaGlyAlaAlaIle 289
Db 66923 GAGATTGTCTCGGTGCTTCCGCGTT 66949
RESULT 12
AAH81345
ID AAH81345 standard; DNA; 864 BP.
XX
AC AAH81345;
XX
DT 21-SEP-2001 (first entry)
XX
DE Escherichia coli protein encoding nucleotide sequence SEQ ID NO:144.
XX
KW Escherichia coli; identification; proliferation; gene therapy; diagnosis;
KW antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;
KW bacterial growth inhibition; ds.
XX
OS Escherichia coli.
XX
PN WC0200148209-A2.
XX

```

```

PD 05-JUL-2001.
XX
XX 19-DEC-2000; 2000WO-US034419.
XX
XX 23-DEC-1999; 99US-017300SP.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Forsyth RA, Ohlsen KL, Zyskind JW;
XX
XX WPI: 2001-457376/49.
XX
XX P-PSDB; AAG98289.
XX
XX Novel nucleic acids encoding proteins required for Escherichia coli
XX proliferation, useful for screening for antimicrobial agents.
XX
XX Claim 9; Page 191-192; 596pp; English.
XX
XX The present invention describes a purified or isolated nucleic acid
XX sequence (I) consisting essentially of one of the 93 nucleotide sequences
XX given in AAH81202 to AAH81294, where expression of the nucleic acid in a
XX microorganism is capable of inhibiting proliferation of a microorganism.
XX (I) have antibacterial and antibiotic activities, and can be used in gene
XX therapy. Expression of (I) in a microorganism inhibits proliferation of
XX the microorganism, and the manufactured antibiotic is useful for reducing
XX the activity or level of a gene product required for proliferation of a
XX microorganism in a subject, specifically humans. The nucleic acids that
XX inhibit bacterial growth or proliferation can be used as antitumor
XX applications, the nucleic acid sequences complementary to sequences
XX required for proliferation can be used as diagnostic tools. For example,
XX nucleic acid probes complementary to proliferation-required sequences
XX that are specific for particular species of microorganisms can be used as
XX probes to identify particular microorganism species in clinical
XX specimens. AAH81295 to AAH81487 encode the Escherichia coli proteins
XX given in AAG98239 to AAG98431, and AAH81488 to AAH81491 represent
XX oligonucleotides, which are used in the exemplification of the present
XX invention
XX
XX SQ Sequence 864 BP; 220 A; 222 C; 240 G; 182 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.6e-97 Length: 864
XX Score: 991.00 Matches: 191
XX Percent Similarity: 82.35% Conservative: 47
XX Best Local Similarity: 66.09% Mismatches: 49
XX Query Match: 69.35% Indels: 2
XX DB: 5 Gaps: 1
XX
XX US-09-545-199F-4 (1-289) x AAH81345 (1-864)
XX
XX Qy 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
XX Db 1 ATGGCCGCGCAAAAGAGATACGTAGTACGAGCGGTCCAGAACCGCAAGATC 60
XX
XX Qy 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
XX Db 61 ACTAAAGCGATGGAGATGGTGGCGCTTCCAAATCGTAAATCGCAGATCGCATGCGC 120
XX
XX Qy 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
XX Db 121 GCCAGCGCTCTTATGACAGAACCATACGACGCGTGTGAGCTGGCGCTTATCGGCTCT 180
XX
XX Qy 61 IleGlyTyrIleHisProPheLeuValAspArgGluValLysValGlyMetIleVal 80
XX Db 181 CTGGAATATAAACATCCGACCTTACCTTGAAGAGCGTGAACCAACCGTGTGCGGTATCGG 240
XX
XX Qy 81 ValSerThrAspArgGlyLeuCysGlyValLeuAsnValAsnLeuPheLysThrValLeu 100
XX Db 241 GTGTGACCGACCGTGGTGTGCGGTGTGTTGACATTAACCTGTTCAAAAGACTGCTG 300
XX
XX Qy 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
XX

```


Db 301 GCGGAATGAAGACCTGCGACGCAAAAGCGGTTCAATGCGACCTCGCAATGATCGGCTCG 360
 Qy 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
 Db 361 AAAGCGGTGGTCTTCACTCGTGGGGCAATGTTGTCACGAGTCCACCGGCAATG 420
 Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
 Db 421 GGGGATAACCCCTTCCTGTCGGAACCTGATCGTCCGGTAAAGTGATGTTGCGAGGCTAC 480
 Qy 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
 Db 481 GACGAAGCCGCTGCGACACCTTATGTCAGCAACAAATTTATTACCAATGCT 540
 Qy 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
 Db 541 CAGGTTCCGACCATCAGCCAGCTGCTCCGTTTACCGGCATCAGATGATGATCTG--- 597
 Qy 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
 Db 598 ---AAACATAAATCTCGGATTACCTGTACGAACCCGATCCGAGCGGTGCTGGATACC 654
 Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
 Db 655 CTGCTGCGTGGTATGTCGAATCTCAGGTTTATCAGGGCGTGGTTGAAAACCTGGCCAGC 714
 Qy 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
 Db 715 GAGCAGCCCGCCGATGTTGGTGGCGATGAAGCCGCGACCGACAATGGCGGCGCTGATT 774
 Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 Db 775 AAAGAGCTGCAGTTGGTATACAAAGCTCGTCAGCCGAGCATTAATCAGGAATCTACC 834
 Qy 281 GluIleValAlaGlyAlaAlaIle 289
 Db 835 GAGATCTCTCGGGGGCCCGCGGTT 861
 RESULT 13
 ACH97650
 ID ACH97650 standard; DNA; 894 BP.
 AC ACH97650;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polynucleotide seqid 3445.
 XX
 KW Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; vaccine; gene; ds.
 XX
 CS Klebsiella pneumoniae.
 XX
 PN US6610836-B1.
 XX
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 PR 29-JAN-1999; 99US-0117747P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL, Osborne M;
 XX
 DR WPI; 2003-895346/92.
 DR F-PDSB; ABO64099.
 XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 PS Disclosure; SEQ ID NO 3445; 932pp; English.
 XX

CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
 CC pneumoniae polypeptide of the invention
 XX

SQ Sequence 894 BP; 223 A; 227 C; 260 G; 184 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.:	4,27e-96	Length:	894
Score:	978.00	Matches:	188
Percent Similarity:	82.35%	Conservative:	50
Best Local Similarity:	65.05%	Mismatches:	49
Query Match:	68.44%	Indels:	2
DB:	11	Gaps:	1

US-09-545-199F-4 (1-289) x ACH97650 (1-894)

Qy	1	MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle	20
Db	31	ATGCGCGGCGCAAAAGAGATACGTAGTAGATCGCAGCGTCCAGACACCGCAAAAGATC	90
Qy	21	ThrLysAlaMetGluMetValAlaLalaSerLysMetArgLysThrGlnGluArgMetSer	40
Db	91	ACTAAAGCGATGGAGATGTCGCCGCTTCCAAAATGCGTAAATCGCAGGCGCATGCGG	150
Qy	41	SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr	60
Db	151	GCCAGCGCTCTTAGCAGATACATATGCGCAAGTATGTCACCTTGGCAAGCGTAAAT	210
Qy	61	IleGlyTyrLysIlePhePheLeuValAspArgGluValLysLysValGlyMetIleVal	80
Db	211	CTGGAATATAAGCACCCCTTACCTGGAAGACGCGAGTTAAAGCGTGGGTACTCTGTG	270
Qy	81	ValSerThrAspArgGlyLeuCysGlyLeuAsnValAsnLeuPheLysThrValLeu	100
Db	271	GTGTGACCGACCGTGTCTGTGCGCGCGCTTGAACATTAACTTGTTCAGAAATCTGTG	330
Qy	101	AsnGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer	120
Db	331	GCGGAATAGAAAGCATGTCGATAAAGCGTTTCAAGTGGACCTCGCAATCATCGGTCT	390
Qy	121	LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle	140
Db	391	AAGGCGTGTCTTCTTTAACTCGGTGGCGGCAACGTTGTCACAGGTGACCGGTATG	450
Qy	141	GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr	160
Db	451	GGGGATAACCCATCCCTGTCGAACTGATCGCCCGGTAAAGTGAATGTCAGGCGCTAT	510
Qy	161	LysLysGlyGluValAspValTyrLeuValTyrAsnLysPheIleAsnThrMetSer	180
Db	511	GATGAAGCCGCTGCGACAGCTGTACGTTGTCAGCAACAAATTTATTAAACCATGTCT	570
Qy	181	GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly	200
Db	571	CAGGTTCCGACGATCACTCAGCTGCTCGCTTACCGGCATCAGAGATGCTGATCTG---	627
Qy	201	GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn	220
Db	628	---AAGCGTAAATCTGGGATTACCTGTATGAGCCCGATCCGAAAGCGCTGCTGGATACC	684
Qy	221	LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer	240
Db	685	CTCTGCGCGCTACGTCGAATCTCAGGTTTATCAGGCGCTGTTGAAAACCTGGCCAGC	744
Qy	241	GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle	260
Db	745	GAGCAGGCGCGCGTATGTTGGCGATGAAAGCGCGACCGATAATGCGCGCAGCTGATT	804
Qy	261	AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn	280


```

Db      805 AAGAGCTCCAGTTGGTATACAAAGCTCGTCAGGCGAGCTACTACGGAATCACC 864
Qy      281 GluileValAlaGlyAlaAlaAlaAla 289
Db      865 GAGATGCTCTCGGGGGCGCGCGGT 891

RESULT 14
ABD03905
ID ABD03905 standard; DNA; 876 BP.
AC ABD03905;
XX
DT 29-JUL-2004 (first entry)
DE Pseudomonas aeruginosa polynucleotide #2509.
KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX
OS Pseudomonas aeruginosa.
PN US6551795-B1.
PD 22-APR-2003.
PF 18-FEB-1999; 99US-00252991.
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR P-PSDB; ABO70334.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 2509; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPIO at
CC seqdata.uspro.gov/sequence.html
XX
SQ Sequence 876 BP; 206 A; 250 C; 266 G; 154 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1-56e-90 Length: 876
Score: 926.50 Matches: 180
Percent Similarity: 79.58% Conservative: 50
Best Local Similarity: 62.28% Mismatches: 56
Query Match: 64.84% Indels: 3
DB: 11 Gaps: 1

US-09-545-199F-4 (1-289) x ABD03905 (1-876)

```

```

Qy      1 MetAlaGlyAlaLysGluileArgThrLysileAlaSerVallySerThrGlnLysile 20
Db      16 ATGGCAGCGCCAAAGAGATTCCGACGAGATTCCGAGCATCAAAAGCAGCGAAAAGATC 75
Qy      21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
Db      76 ACCAATGCCATGGAAAAGTGGCGGTGAGCAAGATGCGCAAGCACAATGCGCATGGCG 135
Qy      41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db      136 GCCGCCGTCCCTACCGGAGCGTATTCCGCAGGTGATCGCCATCTGCGCAACGCGCAAC 195
Qy      61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
Db      196 CCGGAATACCGTCACCGCTTCATGGTCGAGCGTGAAGTCAAGCGCGTCGCTACATCGTG 255
Qy      81 ValSerThrAspArgGlyLeuCysGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db      256 GTGAGCAGTACCGTGTCTGTCGCGCGTCTGAACATCAACCTGTCAAGTCCCTCGTC 315
Qy      101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
Db      316 AAGCAGATGAGCGCTACCGCAGCAGCGGCGGAATCGACCTTTCGCTGCGTAGC 375
Qy      121 LysSerIleAsnPhePheGlnSerLeuGlyIleLysileLeuThrGlnAspSerGlyIle 140
Db      376 AAGGCGCTTCGTTCTTCGCGAGCTTTCGCGCGCAACGTCGTCGAGCTATCAGCCACCTC 435
Qy      141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db      436 GCGAAGAGAGCCTTCGATCAACGACCTGATCGCGAGTGTCAAGGTGATGCTCAGCATAC 495
Qy      161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db      496 CTCGAAGGCGCGTATCGATCGCTGTTCTGCTGCTCCAAACAACTTCGTCAACACCATGACC 555
Qy      181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db      556 CAGAGCCGACCGTGAACAGCTGATTCCGCTGGTGGCGCGATGACGATCAGGAGGTG--- 612
Qy      201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAsn 220
Db      613 -----AAGCACCCTGGGAGTATCTCTACGAACCGCAGCCCAAGTCTCTCTCGACGGG 666
Qy      221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluAsnLeuAlaSer 240
Db      667 CTGCTGTCGCTTACGTGGAATCCCGAGGTGTACCAAGCGCGTGGTGTGAGAACACGCTGT 726
Qy      241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
Db      727 GAGCAGCGCGCGCGGATGATTCAATGAAGAAGCTACCGCAACCGCGCGGAGCTGATC 786
Qy      261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db      787 AGCGATTTCACCTGATCTACACAGCGCGGTGAGCGCGATCAGCCAGGAAATCTCG 846
Qy      281 GluileValAlaGlyAlaAlaAlaAla 289
Db      847 GAAATCGTCGGCGCGCTCGCGCGGTG 873

RESULT 15
ABD03801/C
ID ABD03801 standard; DNA; 1050 BP.
XX
AC ABD03801;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polynucleotide #2405.
XX
KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 9, 2004, 07:03:13 ; Search time 24316 Seconds
(without alignments)
433.093 Million cell updates/sec

Title: US-09-545-199F-4

Perfect score: 1429

Sequence: 1 MAGAKEIRTKASVKSTQKI.....ARQASITNELNEIVAGAAAI 289

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool p/US09545199/runat 08112004 112526 20963/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09545199 -CGN 1 1 6425 -runat 08112004 112526 20963 -NCFU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	693	48.5	699	4	BJ498955
C 2	675	47.2	729	7	CF983748
C 3	605.5	42.4	1469	8	BZ572880
C 4	463	32.4	1437	3	AY108268
C 5	440	30.8	1313	3	CNS0A372
C 6	432.5	30.3	807	8	BH387684
C 7	406	28.4	337	8	BH3817803
C 8	403	28.2	796	8	BH393396
C 9	381	26.7	1474	8	BZ569821

C 10	379.5	26.6	448	8	AQ936809
C 11	352.5	24.7	666	8	BH378165
C 12	350.5	24.5	1205	3	CNS0A385
C 13	349.5	24.5	667	8	BH385025
C 14	342.5	24.0	872	3	CF651710
C 15	340	23.8	1514	3	AY108441
C 16	337.5	23.6	852	7	CN820906
C 17	335	23.4	916	7	CO004140
C 18	335	23.4	917	7	CO008879
C 19	335	23.4	928	7	CO031253
C 20	335	23.4	987	7	CO001829
C 21	334.5	23.4	1064	6	CB686051
C 22	333.5	23.3	838	7	CO364454
C 23	330.5	23.1	358	4	BG876920
C 24	330	23.1	904	5	BU797110
C 25	329	23.0	1036	2	BF942618
C 26	326.5	22.8	858	7	CO088347
C 27	323	22.6	847	4	BG839438
C 28	319.5	22.4	822	7	CK788560
C 29	318.5	22.3	836	7	CO078673
C 30	318	22.3	869	5	EX742945
C 31	318	22.3	872	7	CK191546
C 32	318	22.3	886	7	CK189737
C 33	318	22.3	890	7	CK189738
C 34	318	22.3	898	7	CK189587
C 35	318	22.3	911	7	CK188511
C 36	316	22.1	835	7	CR449718
C 37	316	22.1	876	7	CR414264
C 38	314	22.0	877	7	CO082073
C 39	313	21.9	846	7	CR569311
C 40	313	21.9	1213	3	AY439378
C 41	312.5	21.9	1064	4	BG837916
C 42	312.5	21.9	1083	3	AY375344
C 43	310	21.7	1065	3	AK088164
C 44	310	21.7	1736	3	BC048777
C 45	309	21.6	1104	7	CN061585

ALIGNMENTS

RESULT 1
BJ498955/c
LOCUS
BJ498955 MF01FSA cDNA Oryzias latipes cdna linear EST 08-AUG-2002
mRNA sequence.
VERSION BJ498955.1 GI:22150917
KEYWORDS
SOURCE Oryzias latipes (Japanese medaka)
ORGNISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE
1 (bases 1 to 699)
Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
Unpublished (2001)
COMMENT
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. 699
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="d-r"
/db_xref="taxon:8090"
/clone="MF01FSA015K15"
/sex="mixture of female and male"

FEATURES

source

/tissue type="whole embryo"
/dev stage="fry stage 40"
/clone_lib="XF01P5A CDNA"

ORIGIN

Alignment Scores:

Pred. No.: 2,75e-73 Length: 699
Score: 693.00 Matches: 134
Percent Similarity: 79.33% Conservative: 31
Best Local Similarity: 64.42% Mismatches: 41
Query Match: 48.50% Indels: 2
DB: 4 Gaps: 1

US-09-545-199F-4 (1-289) x BJ498955 (1-699)

QY 82 SerThrAspArgGlyLeuCySGlyGlyLeuAsnValAsnLeuPheLysThrValLeuAsn 101
Db 698 TCGACCGACCGTGTGTGCGGTGTGGAACATTAACTGTTCAAAAACGTCTGGCG 639
QY 102 GLuMetLysGluThrPlySGlyLysAspValSerValGlnLeuSerLeuLeuLeuSerLys 121
Db 638 GAATGAGACCTGGACCGCAAGCGGTCAATGCGACCTCGCAATGCGCTCGAAA 579
QY 122 SerLeuAsnPheGlnSerLeuGlyLysLeuThrGlnAspSerGlyLysGly 141
Db 578 GCGGTGTGCTGTCAACTCGGCGCGCAATGTTGTCACCGAGTCAACGGATGGG 519
QY 142 AspThrProSerValGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 161
Db 518 GATACCTTCCTGTCGCAAGATGTCGTCGCGTAAAGTGAATGTCAGGCTACGAC 459
QY 162 LysGlyGluValAspValValLeuValLeuValLeuValLeuValLeuValLeuVal 181
Db 458 GAAGCGCGTCTGACACAGCTTATACATGTCACCAACAATTTATACACCATGCTCAG 399
QY 182 LysProValLeuGluLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 201
Db 398 GTTCCGACCATCAGCAGCTGTGCGGTGTACCGCATCAGATGATGATCTG----- 345
QY 202 ArgLysGlnValThrAspThrLeuValLeuValLeuValLeuValLeuValLeuVal 221
Db 344 AAACATAAATCTGGATTAATCTGACGACCGGATCGGAGCGGTGTCGATACCCCTG 285
QY 222 LeuValArgThrLeuGluSerGlnValValLeuValLeuValLeuValLeuValLeuVal 241
Db 284 CTGCTGCTGTATGTCGAATCTCAGGTTTATCAGGCGGTGTTGAAACCTGGCCAGCGAG 225
QY 242 GlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAlaValLeuLeuLeuLeuLeu 261
Db 224 CAGCGCGCGGTATGTCGCGTGAAGCCGCGACCAATGCGCGACCTGATATAA 165
QY 262 GluLeuGlnLeuValThrAsnLysAlaArgGlnAlaSerLeuThrAsnGluLeuAsnGlu 281
Db 164 GAGCTGCGAGTGGTATACAAAGAGCTGTCAGGCCAGCATTAATCAGGAACCTCAGGAG 105
QY 282 IleValAlaGlyAlaAlaAlaIle 289
Db 104 ATCGTCTCGGCGCGCGCGCGGT 81

RESULT 2

CF983748/c

LOCUS

DEFINITION

CF983748

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 729)

CF983748 729 bp mRNA linear EST 25-NOV-2003
ID05 grape selected express library Vitis vinifera CDNA 5', mRNA
sequence.

CF983748.1 GI:38516573

Vitis vinifera

Vitis vinifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; Vitaceae; Vitis.

1 (bases 1 to 729)

AUTHORS

TITLE

JOURNAL

COMMENT

Ablett, E., Seaton, G., Scott, K., Shelton, D., Graham, M.W.,
Bavestock, P., Lee, L.S. and Henry, R.
Analysis of grape ESTs: global gene expression patterns in leaf and
berry
Plant Sci. 159 (1), 87-95 (2000)
Contact: Patricio Arce-Johnson
Laboratory of Biochemistry, Department of Molecular Genetics and
Faculty of Biological Sciences, Catholic University of Chile
Alameda 340, Santiago, Chile
Tel: (56-2) 6862897
Fax: (56-2) 2225515
Email: parce@genes.bio.puc.cl
Seq primer: 73

FEATURES

source

1..729
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Chardonnay"
/db_xref="taxon:29760"
/clone_lib="grape selected express library"
/note="Organ: leaf and berry; Vector: pBluescript II SK+;
Site 1: 5' Eco R I; Site 2: 3' Xho I; From leaves total
RNA were isolated as described in Plant Mol Biol Reporter
11 (1993) 212-215. From berries total RNA were isolated
using CTAB method and lithium chloride precipitation
(Plant Mol Biol Reporter 11 (1993) 117-121). Double
stranded cDNA were prepared from 5 ug of polyA leaf RNA or
25 ug total RNA berry RNA using a Stratagene cDNA
synthesis kit. This methods produces hemymethylated cDNAs
cloning in pBluescript SK II+. Following ligation,
plasmids were transformed into XL1-blue-MRF' cells
(Stratagene) and plated for blue-white selection."

ORIGIN

Alignment Scores:

Pred. No.: 4,47e-71 Length: 729
Score: 675.00 Matches: 140
Percent Similarity: 78.38% Conservative: 34
Best Local Similarity: 63.06% Mismatches: 42
Query Match: 47.24% Indels: 6
DB: 7 Gaps: 1

US-09-545-199F-4 (1-289) x CF983748 (1-729)

QY 72 GluValLysLysValGlyMetIleValVal-SerThrAsp-ArgGlyLeuCySGlyGly 91
Db 727 GACGTTAAACGCGTGGGCTACCTGGTGGTTCGACCGCCGCGTGTGTGCGGGTGT 668
QY 91 euAsnValAsnLeuPhe-LysThrValLeuAsnGluMetLysGluThrLysGlyLysAsp 110
Db 667 TGAACATTAACTGTTCAAAAACCTGCTGGCGGAATGAGACCTGGACCGCAAGGC 608
QY 111 ValSerValGlnLeuSer-LeuIleGlySerLysSerIleAsnPheGlnSerLeuGlu 130
Db 607 GTTCAATGCGACCTCGCAAATGATCGGTCGAAAGCGGTGCTCTTCAACTCGGTGG 548
QY 130 YIleLysIleLeuThrGlnAspSerGlyIleGlyAspThrProSerValGluGlnLeuI 150
Db 547 CGCAATGTTGTGGCCAGNTCACCGGCATGGGGATAACCTTCCTGTCGGAACGTAT 488
QY 150 eGlySerValAsnSerMetIleAspAlaThrLysLysGlyGluValAspValValLeu 170
Db 487 CGGTCGGTAAAGTATGTTGTCAGCCCTACGACGAGCGCGTCTGGACAACTTTACAT 428
QY 170 uValThrAsnLysPheIleAsnThrMetSerGlnLysProValLeuGluLysLeuIlePr 190
Db 427 TGCACGACAAATTTATTAACACCATGCTCAGGTTCCGACCATCAGCGCTGCTGCC 368
QY 190 oLeuProGluLeuAspAsnAspGluLeuGlyGluArgLysGlnValThrAspThrLys 210
Db 367 GTTACCGGCATCAGATGATGATCTG-----AAACATAAATCTCGGGATACCTGTA 314

QY 210 rGluProAspAlaLysValLeuLeuAspAsnLeuValArgTyrLeuGluSerGlnVa 230
 Db 313 CGAACCGATCCGAGAGCGTGTCTGGATACCTCCCTGCTGCTGTATGTGATCTCAGGT 254
 QY 230 lTyrGlnAlaAlaValGluAsnLeuAlaSerGluGlnAlaAlaArgMetValAlaMetly 250
 Db 253 TTATCAGGGCGTGTGTGAACCTGCGCAGCAGCAGCGCCCGTATGTGCGCATGAA 194
 QY 250 sAlaAlaThrAspAsnAlaGlyAsnLeuLeuLeuLeuGlnLeuValTyrAsnLysAl 270
 Db 193 AGCCGACCGACAATGGCGCAGCTGATTAAAGAGCTGCGATGTGTATACACAAAGC 134
 QY 270 aArgGlnAlaSerIleThrAsnGluLeuAsnGluIleValAlaGlyAlaAlaIle 289
 Db 133 TCGTCAGCGCAGCATTAATCAGGAATCAACGAGATCGTCTCGGGCGCGCGCGT 76

RESULT 3
 BZ572880 1469 bp DNA linear GSS 17-DEC-2002
 LOCUS
 DEFINITION msh2_2842.x3 msh Pseudomonas aeruginosa genomic clone msh2_2842,
 genomic survey sequence.

ACCESSION BZ572880
 VERSION BZ572880.1 GI:27207941
 KEYWORDS GSS.

SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa

REFERENCE Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.

AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.

TITLE Whole-Genome-Sequence Variation among multiple isolates of

JOURNAL Pseudomonas aeruginosa library

COMMENT J. Bacteriol. (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun

FEATURES

source

Location/Qualifiers

1..1469

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="MSH"

/db_xref="taxon:287"

/clone_lib="msh2_2842"

/note="Environmental isolate. Whole genomic shotgun

library."

ORIGIN

Alignment Scores:

Pred. No.: 3,33e-62 Length: 1469

Score: 505.50 Matches: 120

Percent Similarity: 72.57% Conservative: 44

Best Local Similarity: 53.10% Mismatches: 58

Query Match: 42.37% Indels: 4

DB: 8 Gaps: 1

US-09-545-199f-4 (1-289) x BZ572880 (1-1469)

QY 15 LysSerThrGlnLysIleThrLysAlaMetGluValAlaAlaSerLysMetArgLys 34

Db 43 AGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCGACGATGGCGAAG 102

QY 35 ThrGlnGluArgMetSerSerArgProTyrSerGluThrIleArgAsnValIleSer 54

Db 103 GCACAAATGCGATGGCGCGCGCGCTCCCTACGCGAGCGTATTCGCGAGTATCGCG 162

QY 55 HisValSerLysAlaThrIleGlyTyrLysHisProPheLeuValAspArgGluValLys 74

Db 163 CATCTGGCCAAACCCCAACCCGGAATACCGTCCCGGTTTCATGGTCGAGCTGAAGTCAAG 222
 QY LysValGlyMetIleValValSerThrAspArgGlyLeuCysGlyLeuGlyLeuAsnValAsn 94
 Db 223 CGCGTCGGCTACATCGTGTGTGAGCAGTACCGTGTCTGTGCGCGGCTGTGAACATCAAC 282
 QY LeuPheIysThrValLeuAsnGluMetLysGluIleLysGluLysAspValSerValGln 114
 Db 283 CTCTTCAAGTCCTCGTCAAGGACATGAGCGGCTACCGCGAGCAAGGGCGGAAATCGAC 342
 QY LeuSerLeuIleGlySerLysSerIleAsnPheGlnSerLeuGlyIleLysIleLeu 134
 Db 343 CTTTGGGTGATCGGTAGCAAGGCGCTTCGTTCTTCCCGCAGTTTCGGCGGCAACGTCGTG 402
 QY ThrGlnAspSerGlyIleGlyAspThrProSerValGluGlnLeuIleGlySerValAsn 154
 Db 403 GCAGCTATCAGCACCTCGCGAAGAGCCTTCGATCAACGACCTGATCGGCAGTGTCAAG 462
 QY SerMetIleAspAlaTyrLysLysGlyGluValAspValValTyrLeuValTyrAsnLys 174
 Db 463 GTGATGCTCGACGCATACCTCGAGGCGGTATCGATCGCTGTTCGTGGTCTCCAAACAAG 522
 QY PheIleAsnThrMetSerGlnLysProValLeuGluLysLeuIleProLeuProGluLeu 194
 Db 523 TTCTGTCACACCATGACCCAGAACCGACCGCTGGAACAGCTGATTCGCTGTTGGCCGAT 582
 QY AspAsnAspGluLeuGlyGluArgLysGlnValTyrAspTyrIleTyrGluProAspAla 214
 Db 583 GACGATCAGAGCTG-----AAGCACCATCGGAGCTATCTCTACGAACCGCGACGCC 633
 QY LysValLeuLeuAspAsnLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAl 234
 Db 634 AAGTCGCTCTCGACGGGCTGCTGTTGCGTACGTGGAATCCCAAGTGGACCAAGGCGCGG 693
 QY 234 aValGluAsnLeuAla 239
 Db 694 GGTGAGGACCAACGCC 709

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones,

these are publically available from ZmDB and may be found by BLAST

searching at MSL: maizenap.org; ZmDB, www.zmdb.iastate.edu; TIGR,

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat

Schnable, Iowa State, then clones may be requested from ZmDB:

www.zmdb.iastate.edu.

Location/Qualifiers

1..1437

/organism="Zea mays"

/mol_type="mRNA"

AY108268 1437 bp mRNA linear HTC 16-OCT-2002
 Zea mays PCO078526 mRNA sequence.

AY108268

AY108268.1 GI:21211346

HTC.

Ze mays

Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 1437)

Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,

Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of

Overgo Probes

Unpublished (2002)

2 (bases 1 to 1437)

Coe,E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones,

these are publically available from ZmDB and may be found by BLAST

searching at MSL: maizenap.org; ZmDB, www.zmdb.iastate.edu; TIGR,

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat

Schnable, Iowa State, then clones may be requested from ZmDB:

www.zmdb.iastate.edu.

Location/Qualifiers

1..1437

/organism="Zea mays"

/mol_type="mRNA"

```

/db_xref="MaizeDB:634352"
/db_xref="taxon:4577"
/Clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 6,39e-45 Length: 1437
Score: 463.00 Matches: 114
Percent Similarity: 55.49% Conservative: 63
Best Local Similarity: 35.74% Mismatches: 108
Query Match: 32.40% Indels: 34
DB: 3 Gaps: 5

US-09-545-199F-4 (1-289) x AY108268 (1-1437)

QY 5 LysGluLeuArgThrLysLeuAlaSerValLysSerThrGlnLysIleThrLysAlaMet 24
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 213 CGTGACTCGACCGCGATCGCTCGTCAAGAACACGACAGATCAGGAGCCATG 272

QY 25 GluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSerSerArgPro 44
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 273 AAGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTCTCGCCCT 332

QY 45 TyrSerGluThrIleArgAsnValIleSerHisValSerLys-----AlaThrIleGly 62
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 333 TTCTCGGAGCGCTGTGTGAGTGTCTTACACATGATGAACGAGAGATCCAGACGGAGGAC 392

QY 63 TyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleValValSer 82
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 393 ATCGACTCGCCCTCACCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 452

QY 83 ThrAspArgGlyLeuGlyGlyLeuAsnValLysLysValGlyMetIleValValSer 102
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 453 GCGGAGCGCGCGCTCTCGGAGGCTTCAACACACCGTGTCAAGAGCGGGAACCGCG 512

QY 103 MetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySerLysSer 122
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 513 ATCGATGAGTCAAGACAGCTGGCTCGACGACACCTGCTCGAGCTGGGGAAGAGGCG 572

QY 123 IleAsnPheGlnSerLeuGlyLysLysLysLysLysLysLysLysLysLysLys 142
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 573 AACGCTACTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 632

QY 143 ThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyrLysLys 162
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 633 GTGCCACCGTCAAGGACTCGACGGCCATCGGACCTGCTACTCTACTCTCTCTCTCTCT 692

QY 163 GlyGluValAspValTyrLeuValTyrAsnLysPheIleAsnThrMetSerGlnLys 182
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 693 GAGGAGTGGCAAGGTGGAGTCTCTCTACTCAAGTTCGTGTGCTGTGCGCTCGAC 752

QY 183 ProValLeuGluLysLeuIleProLeuPro----- 192
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 753 CCCATCATCCAGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 812

QY 193 -----GluLeuAspAsnAspGluLeu----- 199
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 813 GTCTGCTGACCGCCACCGAGGAGGATGAGATCGAGACCGACCTTCTCTCTCTCTCTCT 872

QY 200 ---GlyGluArgLysGlnVal-----TrpAspTyrIle-----Tyr 210
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 873 ACCGTGAGCGCGAGAGGTGAGATCGAGACCGACCTTCTCTCTCTCTCTCTCTCTCTCT 932

QY 211 GluProAspAlaLysValLeuLeuAspAsnLeuValArgTyrLeuGluSerGlnVal 230
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 933 GAGCAGACCCCGTGCAGATCTCTGGACGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 992

```

```

QY 231 TyrGlnAlaAlaValGluAsnLeuAlaSerGluGlnAlaAlaArgMetValAlaMetLys 250
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 993 CTCCTGTCCTCAGGAGTGTCTGCCAGCGAGTCTGCCGCCGCGATCAGCGCATGAGC 1052

QY 251 AlaAlaThrAspAsnAlaGlyAsnLeuIleAsnGluLeuGlnLeuValTyrAsnLysAla 270
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1053 AGCGCCACCGACACCGCATCGAGTCTCGCAGAACCTTCTCCATCGCGCTACACCGCAG 1112

QY 271 ArgGlnAlaSerIleThrAsnGluLeuAsnGluLeuValAlaGlyAlaAlaIle 289
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1113 CGCCAGGCCAAGATCACCAGGAGATCTCGAGATCTCGCGCGTGTGCGCGCTC 1169

```

```

CNSOAT3T2 1313 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA complete sequence from clone
GSLTPGH842B07 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).

```

```

ACCESSION BX828395 GI:42461763
VERSION BX828395.1
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

```

```

REFERENCE 1 (Bases 1 to 1313)
AUTHORS Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpetti,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

```

```

JOURNAL Unpublished
REFERENCE 2 (Bases 1 to 1313)
AUTHORS Genoscope.
JOURNAL Direct Submission

```

```

COMMENT Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA, Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full
length

```

FEATURES

```

source
Location/Qualifiers
1..1313
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH842B07"
/tissue_type="Hormone Treated Callus"
/plasmid="pCMVSPORT6"
complement(1..1313)
/gene="At4g04640"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.47e-42 Length: 1313
Score: 440.00 Matches: 113
Percent Similarity: 53.58% Conservative: 59
Best Local Similarity: 35.20% Mismatches: 115
Query Match: 30.79% Indels: 34

```

DB: 3 Gaps: 5

US-09-545-199F-4 (1-289) x CNS0A3T2 (1-1313)

Qy 2 AlaGlyAlaLysGluLeuThrLysIleAlaSerValLysSerThrGlnLysIleThr 21
 Db 128 GCCTCTCTCGTGGAGCTCAGACCGTATCGATTGAGTCAAGGACACTCAAGATCACC 187

Qy 22 LysAlaMetGluMetValAlaLysSerLysMetArgLysThrGlnGluArgMetSerSer 41
 Db 188 GAAGCTATGAAGCTGTCTGCTCAGCTAAAGTCAGGAGCTCAGAGCTGTGTCTCAAT 247

Qy 42 SerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLys-----Ala 59
 Db 248 GCACGACCATCTCTGAAACCTTACTGAAGTCTTTTCAACATCAACGACGACGTTCAA 307

Qy 60 ThrIleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIle 79
 Db 308 ACCGATGATGTCGATGTTCTTAAACAAAGTCAGACCGTTAAGAAGTAGCTCTCGTT 367

Qy 80 ValValSerThrAspArgGlyLeuGlyCysGlyLeuAsnValLysValLysThrVal 99
 Db 368 GTCGTCACCGGTATCGTGGATTATGTTGGTATTCACCAATTCATCAATTAAGAAGCA 427

Qy 100 LeuAsnGluMetLysGluThrLysGluLysAspValSerValGlnLeuSerLeuIleGly 119
 Db 428 CAGGCAAGATCAAGAGCTTAAAGGCTAGGCTTGAATACACAGTCAATAGCGTGGC 487

Qy 120 SerLysSerIleAsnPhePheGlnSerLeuGlyLysIleLeuThrGlnAspSerGly 139
 Db 488 AAGAAGGAATTTCTTCTCCGTCGCGCTACATCCCGTCACAAATACCTAGAA 547

Qy 140 IleGlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspIle 159
 Db 548 CCGGGAATTTACTACGGCTAAGAAAGCTCAAGCTGTCGCTGATGATGCTCTCTCTG 607

Qy 160 TyrLysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMet 179
 Db 608 TTTATAGTGAAGATCGAAAGTCGAGCTCTGTACACAAAGTTGTATCTTGTGTC 667

Qy 180 SerGlnLysProValLeuGluLysLeuIleProLeuPro----- 192
 Db 668 AATCAGAACCCGTGATCTACAGCTACTGCTGCTTTATCACCTAAAGGAGATCTGTGAC 727

Qy 193 -----GluLeuAspAsnAspGluLeu----- 199
 Db 728 ATTAATGGAACCTGTGTGGATGCTCGGAAGATGAATTTTCAGTTAAACGACAAAGAA 787

Qy 200 GlyGlu-----ArgLysGlnValTyrAspTyrIle----- 209
 Db 788 GGGAAATTCAGACTTGAAGAGAGACTTTTAGGACACCAACAGCTGATTTCTCGCCGATC 847

Qy 210 -----TyrGluProAspAlaLysValLeuLeuAspAsnLeuLeuValArgTyrLeuGlu 227
 Db 848 TTGCAATTCGAGCAAGACCTGTTCAGATCTTCTGATCTTGTGCTCTGTATCTTAAC 907

Qy 228 SerGlnValTyrGlnAlaLysValGluAsnLeuLysSerGluGlnAlaLysMetVal 247
 Db 908 AGTCAGATCTTAGGCGATACAGAGTCACTGTGCTAGTGTGAGCTTGCAGTAGAATGAT 267

Qy 248 AlaMetLysAlaLysThrAspAsnAlaGlyAsnLeuIleAsnGluLeuGlnLeuValTyr 267
 Db 968 GCAATGAGTAGTCTCGGATAATGCATCGATCTCAAGAAATCGCTTTCGATGGTGTAT 1027

Qy 268 AsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsnGluIleValAlaGlyAlaIle 287
 Db 1028 AATGAAGCGTCAAGCTAAGATTACTGGAGAGATTTCTTGAGATTTCTGCGAGCTAAT 1087

Qy 288 Ala 288
 Db 1088 GCA 1090

RESULT 6

BH387684 807 bp DNA linear GSS 11-DEC-2001
 LOCUS AG-ND-167A14.TR ND-TAM Anopheles gambiae genomic clone
 DEFINITION AG-ND-167A14, genomic survey sequence.
 ACCESSION BH387684
 VERSION BH387684.1 GI:17333825
 KEYWORDS GSS.
 SOURCE Anopheles gambiae (African malaria mosquito)
 ORGANISM Anopheles gambiae
 Eukaryota; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles
 1 (bases 1 to 807)
 REFERENCE Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J., Ren,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B., Gardner,M.J. and Collins,F.H.
 TITLE Construction of a BAC library and generation of BAC end sequence-tagged connectors for genome sequencing of the African malaria mosquito Anopheles gambiae
 JOURNAL Mol. Genet. Genomics 268 (6), 720-728 (2003)
 MEDLINE 22542063
 PUBMED 12655398
 COMMENT Other GSSs: AG-ND-167A14.TF
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjloftus@tigr.org
 This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.
 Seq primer: M13 Rev
 Class: BAC ends.
 FEATURES
 Location/Qualifiers
 1..807
 /organism="Anopheles gambiae"
 /mol_type="genomic DNA"
 /strain="PEST"
 /db_xref="taxon:7165"
 /clone="AG-ND-167A14"
 /clone_lib="ND-TAM"
 /note="Vector: pECBAC1; Site 1: HindIII"

ORIGIN

Alignment Scores:
 Pred. No.: 1,41e-41 Length: 807
 Score: 432.50 Matches: 102
 Percent Similarity: 62.13% Conservative: 67
 Best Local Similarity: 37.50% Mismatches: 95
 Query Match: 30.27% Indels: 8
 DB: 8 Gaps: 4

US-09-545-199F-4 (1-289) x BH387684 (1-807)

Qy 19 LysIleThrLysAlaMetGluMetValAlaLysSerLysMetArgLysThrGlnGluArg 38
 Db 803 CAATACAGTCTATGAAATGTTTCGCGAGCAACCTTAAGAAAGCAGGATGCT 744

Qy 39 MetSerSerSerArgProTyrSerGluThrIleArgAsnValIle-SerHisValSerLys 58
 Db 743 ATTGTGATGCTTCGCCCTTACTCCGAAAACTTCAGGAATAATAGAAAAACGTTAGT-- 686

Qy 58 sAlaThrIleGlyTyrLysHisProPheLeuValAsp-----ArgGluValLysLys 75
 Db 685 -TCTACTAGATCAGGAGACGCTGCTGTGTTGATGTCACACACGAGGAGGTTAAAG 627


```

QY 75 sValGlyMetileValSerThrAspArgGlyLeuCyGlyClyLeuAsnValAsnLe 95
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
626 ATCTGTCATTCAGTACTTCAACAGAGGCTTCGAGGAGCTTCAACTCTCTGT 567
QY 95 uPheIysThrValLeuAsnGluMetLysGluTrpLysGluLysAspValSerValGlnLe 115
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
566 TATCAAGAAGCTTAATGCTCAGTATCAGCAA-----AATTCAGCTTTTTCAGGTTGAAGT 513
QY 115 uSerLeuIleGlySerLysSerIleAsnPheGlnSerLeuGlyLysIleValLeuTh 135
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
512 GTTGACTATGTTAAAGAAAGCATTTGATGCTTTTCAGAGCT---TCTAAGAAATCTATGA 456
QY 135 rGlnAspSerGlyIleGlyAspThrProSerValGluGlnLeuIleGlySerValAsnSe 155
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
455 TAACCACAGGCGATCTTTACGATCTTTAACTTTTGAAGCGGTAGCACATGTAGCTGAGGA 396
QY 155 rMetIleAspAlaTrpLysGlyGluValAspValValTrpLeuValTrpAsnLysPhe 175
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
395 TATCATGAGAGATTTCCGTCAGGAGGTTTGTGATGAAGTATACCTGGTGTACACAAAT 336
QY 175 eIleAsnThrMetSerGlnLysProValLeuGluLysLeuIleProLeuProGluLeuAs 195
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
335 CCTTAATCGCGCAACGACGAGAGTACAGACAGAAAGCTTCTCTATTACATCCCTGA 276
QY 195 pAsnAspGluLeuGlyGluArgLysGlnValTrpAspTrpIleTyrGluProAspAlaLy 215
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
275 AAAGAAGAAGCTGAGAGAGTGCATCGAAACCGATTATATCTTCGAACCAACCCGTAC 216
QY 215 sValLeuLeuAspAsnLeuValArgTyrLeuGluSerGlnValTrpGlnAlaAlaVala 235
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
215 CGAAATTCGGAACCTTAATCTCTAAATCTATCAAGACTCAGGTTTATAAGGATCCCT 156
QY 235 lGluAsnLeuAlaSerGluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAs 255
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
155 TGATTCAAGAGCTTCTGAGCATGGTCAAGAGTACTGCAATGCACAAAGCAACAGATAA 96
QY 255 nAlaGlyAsnLeuIleAsnGluLeuGlnValTrpAsnLysAlaArgGlnAlaSerIl 275
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
95 TGCACAGGCGATTCGAGAGTACCTGGTATCTTCTACACAAAGCACGCTCAGGCTGCAAT 36
QY 275 eThrAsnGluLeuAsnGluIleValAlaGlyAla 286
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
35 TACCAACGAGATCTTGGAAATTTGTATCCGAGCA 2

```

```

RESULT 7
BH817803 337 bp DNA linear GSS 20-MAY-2002
LOCUS BACP10-G19.z Pristionchus pacificus BAC ends Pristionchus
DEFINITION pacificus genomic, genomic survey sequence.
ACCESSION BH817803
VERSION BH817803.1 GI:20989377
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
AUTHORS 1 (bases 1 to 337)
Srinivasan,J., Sinz,W., Lanz,C., Brand,A., Nandakumar,R.,
Raddatz,G., Witte,H., Keller,H., Kipping,I., Pires da Silva,A.,
Jesse,T., Millare,J., de Both,M., Schuster,S.C. and Sommer,R.J.
TITLE A BAC-based genetic linkage map of the nematode Pristionchus
JOURNAL pacificus
COMMENT Unpublished (2002)
CONTACT: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
Location/Qualifiers

```

```

source 1..337
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="var. California"
/db_xref="taxon:54126"
/clone_lib="Pristionchus pacificus BAC ends"
ORIGIN
Alignment Scores: 6.62e-39 Length: 337
Pred. No.: 405.00 Matches: 82
Score: 82.30% Conservative: 11
Percent Similarity: 82.30% Mismatches: 18
Best Local Similarity: 72.57% Indels: 2
Query Match: 28.41% Gaps: 1
DB: 8
US-09-545-199F-4 (1-289) x BH817803 (1-337)
QY 170 LeuValTyrAsnLysPheIleAsnThrMetSerGlnLysProValLeuGluLysLeuIle 189
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3 ATTGTGAGCAACAATTTATTAAACACCATGTTCTCAGGTTCCGACCATCAGCCAGCTGCTG 62
QY 190 ProLeuProGluLeuAspAsnAspGluLeuGlyGluArgLysGlnValTrpAspTrpIle 209
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 CCGTTACCGGCATCAGATGATGATCTG-----AAACATTAATCCTGGGATTAACCTG 116
QY 210 TyrGluProAspAlaLysValLeuLeuAspAsnLeuValArgTyrLeuGluSerGln 229
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 TACGAACCCGATCCGAGCGCTGCTGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 176
QY 230 ValTyrGlnAlaAlaValGluAsnLeuAlaSerGluGlnAlaAlaArgMetValAlaMet 249
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 GTTTATCAGGCGGTGTTGAACCTGGCCAGCAGGAGCCGCCCTGATGGTGGCGCATG 236
QY 250 LysAlaAlaThrAspAsnAlaGlyAsnGluLeuIleAsnGluLeuGlnValTyrAsnLys 269
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 AAAGCCGCGACCGCAATATGGCGAGCGCTGATTAAGAGCTGCAGTTGGTATACACAA 296
QY 270 AlaArgGlnAlaSerIleThrAsnGluLeuAsnGluIle 282
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
297 GCTGTCAGGCGCAGCATTAATCTCAGAACTACCGAGATC 335
RESULT 8
BH393596/6 796 bp DNA linear GSS 11-DEC-2001
LOCUS AG-ND-142J4.TR ND-TAM Anopheles gambiae genomic clone AG-ND-142J4,
DEFINITION genomic survey sequence.
ACCESSION BH393596
VERSION BH393596.1 GI:17339737
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
REFERENCE Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
AUTHORS Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
1 (bases 1 to 796)
Hong,Y.S., Hogen,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J.,
Ren,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B.,
Gardner,M.J. and Collins,F.H.
TITLE Construction of a BAC library and generation of BAC end
sequence-tagged connectors for genome sequencing of the African
malaria mosquito Anopheles gambiae
JOURNAL Mol. Genet. Genomics 268 (6), 720-728 (2003)
COMMENT MEDLINE 22542063
PUBMED 12655398
OTHER_GSSs: AG-ND-142J4.TF
CONTACT: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org

```



```

Db 312 GCCGGCCGTCCTACCGGAGCGTATTTCGCCAGGTGATCGGCATCTGCCCAACGCCAAC 253
Qy 61 lledgtyrlyshisPropheLeuValAspArgGluVallyshysValGlyMetileVal 80
Db 252 CCGGAATACCGTACCCCGTTCATGTCGAGCGTGAAGTCAAGCGCTCGCTACATCGTG 193
Qy 81 ValSerThrAspArgGlyLeuCyGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db 192 GTGAGCAGTACCGCTGCTGTCGCGCGGTCTGACATCATCACTGTTCAAGTCCCTCGTC 133
Qy 101 AsnGluMetLysGluTrpLysGluLys 109
Db 132 AAGGACATGAGCGCTACCGCGCA 106

RESULT 10
LOCUS AQ936809/c
DEFINITION N81-026R Human NotI clones:Homo sapiens genomic, genomic survey
sequence.
ACCESSION AQ936809
VERSION AQ936809.1 GI:7213187
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 448)
Zabarovskiy,E.R., Gizatullin,R., Podowski,R.M., Zabarovska,V.,
Xie,L., Muravenko,O.V., Kozirev,S., Petrenko,L., Skobeleva,N.,
Li,J., Protocopov,A., Kashuba,V., Ernberg,I., Winberg,G. and
Wahlstedt,C.
TITLE NotI clones in the analysis of the human genome
JOURNAL Nucleic Acids Res. 28 (7), 1635-1639 (2000)
MEDLINE 20175728
PUBMED 10710430
COMMENT Contact: Podowski RM
Center for Genomics Research
Karolinska Institute
17177 Stockholm, Sweden
Tel: +46-8-728-6372
Fax: +46-8-337583
Email: Raf.Podowski@cgr.ki.se
Class: NotI site.
Location/Qualifiers
1..448
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human NotI clones"

ORIGIN
Alignment Scores:
Pred. NO.: 1.64e-35 Length: 448
Score: 379.50 Matches: 85
Percent Similarity: 70.27% Conservative: 19
Best Local Similarity: 57.43% Mismatches: 37
Query Match: 26.56% Indels: 7
DB: 8 Gaps: 2

US-09-545-199f-4 (1-289) x AQ936809 (1-448)

Qy 144 ProSerValGluInLeuLleGlySerValAsnSerMetileAspala-TyrLysLysG 163
Db 433 CCCACATCGAGGCCGCTGATCGGCCCGCTCAAGTGTCTGTCGAGCCNCTACGCCAAGG 374
Qy 163 yGluValAspValValTyrLeuValTyrAsnLysPheileAsnThrMetSerGlnLysPr 183
Db 373 CAGGCTGTGCGCGTCTACTCTGCTACACCAAGTTCATCACAGATGAAGCAAGAGGT 314
Qy 183 oValLeuGluLys-LeuileProGluLeuAspAsnAspGluLeuGlyGluArgL 203
Db 313 GCAGGTGCACCGACTCTGCGCTCAAGGCC-----AGCGAGTGGCGCGGAA 263

```

```

Qy 203 ysgInVal-----TTPAspTYrlleTYrGluProAspAlalysValLeuLeuAspAsnL 221
Db 262 AGCGGCCCCACTCGTGGGATTACATCTACGAGCCCGATGCCNCGACNGTCATCATCGAAC 203
Qy 221 euleuValArgTyrLeuGluSerGlnValTyrGlnAlaValAlaValGluAsnLeuLeuAsrG 241
Db 202 TGTGTGTGGCTACTCGGAAGCGTGTCTTCCAGGCGCTGCCGAGAACATGCCAGCG 143
Qy 241 luGlnAlaAlaArgMetValAlaValTyrAsnLysAlaAlaThrAspAsnAlaGlyAsnLeuLea 261
Db 142 AGCAGTCCCGCCGCTGTGGCCATGAAGGCTGGACCGCNCACCGCGCGCGCTGATCG 83
Qy 261 snGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsnG 281
Db 82 GCGAGCTGAAGCTGTCTACANCAAGACCCGCCAGCGGCCCATCCCAAGAGTTGTCCG 23
Qy 281 luileValAlaGlyAlaAla 287
Db 22 AGATGCTCAGCGCGCGCGCC 3

```

```

RESULT 11
LOCUS BH378165/c
DEFINITION

```

```

AG-ND-138P6.TF ND-TAM Anopheles gambiae genomic clone AG-ND-138P6,
genomic survey sequence.
ACCESSION BH378165
VERSION BH378165.1 GI:17324307
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae

```

```

REFERENCE 1 Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
AUTHORS Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
1 (bases 1 to 666)
Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J.,
Ren,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B.,
Gardner,M.J. and Collins,F.H.
TITLE Construction of a BAC library and generation of BAC end
sequence-tagged connectors for genome sequencing of the African
malaria mosquito Anopheles gambiae
JOURNAL Mol. Genet. Genomics 268 (6), 720-728 (2003)
MEDLINE 22542063
PUBMED 12655398
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org

```

```

This clone is from an A. gambiae BAC library (ND-TAM) provided by
F.H. Collins and sequenced by The Institute for Genomic Research
(TIGR). The BAC library was generated from A. gambiae PEST strain
DNA. All DNA was extracted from newly hatched first instar larvae
to minimize the inclusion of DNA from microorganisms that inhabit
the gut. The DNA is derived from mixed sexes of larvae. The BAC
library was constructed at Texas A&M University BAC Center
University, College Station, Texas 77843-2123, USA using a HindIII
partial digest.
Seq primer: M13 For
Class: BAC ends.
Location/Qualifiers
1..666
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strains="PEST"
/db_xref="taxon:7165"
/clone_lib="AG-ND-138P6"
/notes="Vector: pECBAC1; Site_1: HindIII"

```

```

FEATURES
source

```

```

ORIGIN

```

```

Alignment Scores:

```

```

Pred. No.: 5.48e-32 Length: 666
Score: 352.50 Matches: 80
Percent Similarity: 61.19% Conservative: 54
Best Local Similarity: 36.53% Mismatches: 82
Query Match: 24.67% Indels: 3
DB: 8 Gaps: 2

US-09-545-199F-4 (1-289) x BH378165 (1-666)

QY 70 AspArgGluValLysLysValGlyMetIleValValSerThrAspArgGlyLeuGly 89
Db 648 GAGAGTAGGTAAAGAAATCCCTGTCATTCGAGTGCATCTCCACAGAGGTTCGAGGA 589
QY 90 GlyLeuAsnValLeuLeuPheLysThrValLeuAsnGluMetLysGluTrpLysGluLys 109
Db 588 GCTTTCAACTCTCTGTATTATCAAGAAGTAAATGCTCAGTATCAGCAA-----AATCA 535
QY 110 AspValSerValGlnLeuSerLeuIleGlySerLysSerIleAsnPhePheGlnSerLeu 129
Db 534 GCITTTGAGGTTGAGGTGACATATTGGTAACACAGCAATTTGATGCTTTCAGAGCT--- 478
QY 130 GlyLeuLysIleLeuThrGlnAspSerGlyIleGlyAspThrProSerValGluGlnLeu 149
Db 477 TCTAAGACATCTATGATAACACAGCATCTTACGATCTTTAACTTTGAAGCGTA 418
QY 150 IleGlySerValAsnSerMetIleAspAlaTyrLysLysGlyGluValAspValTyr 169
Db 417 GCACATGTAGTCAGATATCATGAGAGATTTCCGTCAGGGAGTTTGTGATGAAGTATAC 358
QY 170 LeuValTyrAsnLysPheIleAsnThrMetSerGlnLysProValLeuGluLysLeuIle 189
Db 357 CTGGTGTACAAACAATTCCTTAATCGGCAACGACGAGGAGTACACAGAGAAACCTTCTT 298
QY 190 ProLeuProGluLeuAspAsnAspGluLeuGlyGluArgLysGlnValTrpAspTyrIle 209
Db 297 CCTATTACATGCTCTCAAGAAAGAAAGCTGAGAGAGTGCATCGAACCGATTAATC 238
QY 210 TyrGluProAspAlaLysValLeuLeuAspAsnLeuValArgTyrLeuGluSerGln 229
Db 237 TTGCAACCAACCGTACCGAATCTGGAACCTTAATCTCTAATCTATCAAGACTCAG 178
QY 230 ValTyrGlnAlaValGluAsnLeuLysGlnAlaAlaArgMetValAlaMet 249
Db 177 GTTTATAGCGATCTCTGATTTCAGTCTCTGAGAAAGAAAGCAAGGATGCTGCAATG 118
QY 250 LysAlaAlaThrAspAsnAlaGlyAsnLeuIleAsnGluLeuGlnLeuValTyrAsnLys 269
Db 117 CACAAGCAACAGATAATGCACAGCATTTGAAGATGACCTGCTGTATCTTCTACACAA 58
QY 270 AlaArgGlnAlaSerIleThrAsnGluLeuAsnGluIleValAlaGlyAlaAla 288
Db 57 GCACGTGAGGCTGCAATTACCAACGAGATCTTGAAATTTGATCCGAGCAGAGACT 1

RESULT 12
LOCUS CNS0A9BS 1205 bp mRNA linear HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSUTLS16ZE03 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
BX819675
GI:42468345
VERSION HTC; GSLT cDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1205)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome

```

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Annotation
Unpublished
2 (bases 1 to 1205)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source

1..1205
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSUTLS16ZE03"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT 6"
complement(1..1205)
/gene="At2g33040"

gene

ORIGIN

Alignment Scores:

Prod. No.: 2.24e-31 Length: 1205
Score: 350.50 Matches: 106
Percent Similarity: 49.84% Conservative: 50
Best Local Similarity: 33.87% Mismatches: 88
Query Match: 24.53% Indels: 69
DB: 3 Gaps: 12

US-09-545-199F-4 (1-289) x CNS0A9BS (1-1205)

```

QY 7 IleArgThrLysIleAlaSerValLysSerThrGlnLysIleThrLysAlaMetGluMet 26
Db 245 GTGCGTAACCGCATGAGAGTGTTAAGACATCAAAAGATCAAAAGGCAATGAAGATG 304
QY 27 ValAlaAlaSerLysMetArgLysThrGlnLysArgMetSerSerArgProTyrSer 46
Db 305 GTTCTGCTTCCAGCTTAGAGCAGTTCAGAGCGGCGAGCTGAGAACTCCCGT----- 355
QY 47 GluThrIleArgAsnValIleSerHisValSerLysAlaThrIleGlyTyrLysHisPro 66
Db 356 -----LeuValAspArg-----GluValLysValGlyMetIleVal 370
QY 67 Phe-----LeuValAspArg-----GluValLysValGlyMetIleVal 80
Db 371 TTACTGCATCTTAGGAGATATCCAGCATTTAGTAAAGAGAGTGTGTGTGTCAC 430
QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThr----- 98
Db 431 CTCTCTTGTGACAAGGGTCTCTGTGTGTGAATCACTCCACTGTCTGTTAAAGTAGCAGG 490
QY 99 ValLeuAsnGluMetLysGluTrpLysGlyLysAspValSerValGlnLeuSerLeuIle 118
Db 491 GCTCTGTACAAATTTGAATGCTGCTCTCTGAAAGGAA-----GTTCACTTTGTTATGTC 544
QY 119 GlySerLysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSer 138
Db 545 GGAGAGAAAGCA-----AAGGCTATAATGTTTCGTGACTCA 580
QY 139 GlyIleGlyAspThrProSerValGluGlnLeu-----IleGlySer 152

```

```

Db 581 AAGAACGACATGTGCTCTCTGTAACAGAGCTGAATAAGAACCCACTCAATTATGCTCAG 640
Qy 153 ValAsnSerMetIle---AspAlaTyrLysGlyGluValAspValValTyrLeuVal 171
Db 641 GTGTCACTTCTAGCTGATGATCCCTGGAAGACGTTGAATTGATGCTTTGCGCATTTGC 700
Qy 172 TyrAsnLysPhe-----IleAsnThrMetSerGlnLys 182
Db 701 TACAACAAGTTCATTTCAGTTGCGCATTTCTGCCAACTGTGCTCCTGCTTTTGCACCT 760
Qy 183 ProValLeuGluLysLeuLe---ProLeuProGluLeuAspAsnGlu 198
Db 761 GAGATTATTGAGAGAGAGCTGGAATTGGAGAAACCTTGGCGAGCTTGAATCATATGAG 820
Qy 199 Leu-----GlyGluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysVal 216
Db 821 ATTGAAGTGGGGGAACAAG-----GGAGAA 847
Qy 217 LeuLeuAspAsnLeuValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGlu 236
Db 848 ATATTGCAAGATCTGGCGAGTCCAAATCTCTGTGTGATGTTCAATGCGGTCTCGAG 907
Qy 237 AsnLeuAlaSerGluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAla 256
Db 908 AATGCATGTAGTGAGATGGAGCAAGATGTCGCCATGGACAGCTCAACGAAAGCGCA 967
Qy 257 GlyAsnLeuLeuAsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThr 276
Db 968 GGAGAAATGCTCGACCGCTCACCTCATACACAGGAGCTGCTCAAGCTTCTATTACA 1027
Qy 277 AsnGluLeuAsnGluLeuValAlaGlyAlaAlaLys 289
Db 1028 ACAGAGCTTATTGAGATTCTCTGGAGCTTCTGCTCTT 1066

RESULT 13
LOCUS BH385025/5
DEFINITION AG-ND-167A6, TR ND-TAM Anopheles gambiae genomic clone AG-ND-167A6,
ACCESSION BH385025
VERSION BH385025.1 GI:17331167
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
1 (bases 1 to 667)
Hong, Y. S., Hogan, J. R., Wang, X., Sarkar, A., Sim, C., Loftus, B. J.,
Rep, C., Huff, E. R., Carlisle, J. L., Black, K., Zhang, H. B.,
Gardner, M. J., and Collins, F. H.
Construction of a BAC library and generation of BAC end
sequence-tagged connectors for genome sequencing of the African
malaria mosquito Anopheles gambiae
Mol. Genet. Genomics 268 (6), 720-728 (2003)
22542063
12655398
Other_GSSs: AG-ND-167A6.TF
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by
F.H. Collins and sequenced by The Institute for Genomic Research
(TIGR). The BAC library was generated from A. gambiae PEST strain
DNA. All DNA was extracted from newly hatched first instar larvae
to minimize the inclusion of DNA from microorganisms that inhabit
the gut. The DNA is derived from mixed sexes of larvae. The BAC
library was constructed at Texas A&M University BAC Center

```

```

University, College Station, Texas 77843-2123, USA using a HindIII
partial digest.
Seq primer: M13 Rev
Class: BAC ends.
FEATURES
    Location/Qualifiers
        1..667
            /organism="Anopheles gambiae"
            /mol_type="genomic DNA"
            /strains="PEST"
            /db_xref="taxon:7165"
            /clone="AG-ND-167A6"
            /clone_lib="ND-TAM"
            /note="Vector: pECBAC1; Site_1: HindIII"
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1.27e-31 Length: 667
Score: 349.50 Matches: 85
Percent Similarity: 60.71% Conservative: 51
Best Local Similarity: 37.95% Mismatches: 81
Query Match: 24.46% Indels: 8
DB: Gaps: 3

US-09-545-199F-4 (1-289) x BH385025 (1-667)
Qy 63 TyrLysHisProPheLeuValAspArgGluValLysValGlyMetIleValValSer 82
Db 651 TATCCACAACA-----AGAGAGGTAAGAAATCTGTTCATTGCGAGTTACT 604
Qy 83 ThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeuAsnGlu 102
Db 603 TCCAAACAGAGGCTTCGAGGAGCTTCAACTCTCTCGTTATCAAGAAGCTTATGCTCAG 544
Qy 103 MetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySerLysSer 122
Db 543 TATCAGCAA-----AATTCAGCTTTTGGGTGAAGTTGCTGACTATTGGTAAAAAGCA 490
Qy 123 IleAsnPhePheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIleGlyAsp 142
Db 489 TTTGATGCTTTTCAGAGCT--TCTAAGAAATCTATGATAACCAACGAGGATCTTACGAT 433
Qy 143 ThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyrLysLys 162
Db 432 TCTTTAACTTTTGAAGCGGTAGACATGATGAGGATATCATGAGAGATTTCGTCAG 373
Qy 163 GlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSerGlnLys 182
Db 372 GGAGTTTTTGTGATGAAGTATACCTGGGTACACAAATTCCTTAAATCGCGCAACGAGGAA 313
Qy 183 ProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGlyLysArg 202
Db 312 GTACAGACAGAAAGCTTCTTCTTATTACATTCGCTGAAAGAAAGAGCTGAGAGAGT 253
Qy 203 LysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsnLeuLeu 222
Db 252 GCAATCGAAACCGATTATATCTTCGAAACCAACCGTCCCGAAATTCGGAACCTTAAT 193
Qy 223 ValArgTyrLeuGluSerGlnValTyrGlnAlaAlaValGluAsnLeuAlaSerGluGln 242
Db 192 CCTAAATCTATCAAGACTCAGGTTTATAAAGCGATCTCTGATTTCAGTAGCTTCTGAGCAT 133
Qy 243 AlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIleAsnGlu 262
Db 132 GGTGCAAGGATGACTGCAATGCACAAGCAACAGATAATGCACAGGATTGGAAGATGAC 73
Qy 263 LeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGlnLeuLeuAsnGluLe 282
Db 72 CTGGTTATCTTCTCAACAAGACGCTCAGGTCGATTCGATTCACCAACGAGATCTT-GAAATT 14
Qy 283 ValAlaGlyAla 286
Db 13 GTATCCGAGCA 2

```

QY	30	SerLysMetArgLysThrGlnGlnArgMetSerSerSerArgProTyrSerGluThrIle	49
Db	63	TCCAAGCTTAGACGAGTTCAGGCGGAGCTGAGAACTCCCGT	104
QY	50	ArgAsnValIleSerHisValSerLysAlaThrIleGlyTyrLysHisProPhe	67
Db	105	-----GGACTTTGGCAGCCATTACTGCA	128
QY	68	---LeuValAspArg-----GluValLysLysValGlyMetIleValValSerThr	83
Db	129	CTTCTAGGAGATAATCCAGCATTGATGTAAGAGAGAGTGTGGTCTCTCTCTCT	188
QY	84	AspArgGlyLeuCysGlyClyLeuAsnValAsnLeuPheLysThr-----ValLeuAsn	101
Db	169	GACAGGGTCTCTGTGGTGGAAATCACTCCACTCTGCTTAAAGTGAGCAGGCGCTGTGAC	248
QY	102	GluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySerLys	121
Db	249	AAATTGAATGCTGGTCTCTGAAAGAA-----GTTCACTTTGTTATTCGCGAGAGAA	302
QY	122	SerIleAsnPhePheGlnSerLeuGlyLleLysIleLeuThrGlnAspSerGlyIleGly	141
Db	303	GCA-----AAGGCTATAATGTTTCGTGACTCAAGAAGACAC	338
QY	142	AspThrProSerValGluGlnLeu-----IleGlySerValAsnSer	155
Db	339	ATTGTCTCTCTGTAAACAGAGCTGAATAAGAACCCACTCAATTATGCTCAGGTGTCAGTT	398
QY	156	MetIle---AspAlaTyrLysLysGlyGluValAspValValTyrLeuValTyrAsnLys	174
Db	399	CTAGCTGATGACATCTCGAAGAACGTTGAATTGATGCTTTCGCGATTTGCTACACAAG	458
QY	175	Phe-----IleAsnThrMetSerGlnLysProValLeu	185
Db	459	TTTCATTCACTGTGCGATTCTGCGCAACTGTGTCCTGTTTCTCAGCTGAGATTATT	518
QY	186	GluLysLeuIle-----ProLeuProGluLeuAspAsnAspGluLeu-----	199
Db	519	GAGAAGGAGTCTGAATTGGAGAAACTTTGGCGAGCTTGACTCATATGAGATTGAAAGGT	578
QY	200	GlyGluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAsp	219
Db	579	GGGGAAACAAAG-----GGAGAAATATTGCAG	605
QY	220	AsnLeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAla	239
Db	606	AATCTGCGCAGTTCCAAATCTCTTGTGTGATGTTCAATGCGGTCTCGAGAATGCATCT	665
QY	240	SerGluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeu	259
Db	666	AGTGAGATGGGAGCAAGAATGTCCTCCATGGACAGCTCAAGCAGAAACGCGAGAGAAATG	725
QY	260	IleAsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeu	279
Db	726	CTCGACCGTCTACCCCTCACATACACAGGAGCTCGTCAAGCTTCTATTACACAGAGCTT	785
QY	280	AsnGluIleValAlaGlyAlaAlaAlaIle	289
Db	786	ATTGAGATTATCTCTGGAGCTTCGCTCTT	815
RESULT 15			
LOCUS			
AY108441			
DEFINITION			
Zea mays P0083558 mRNA sequence.			
ACCESSION			
AY108441			
VERSION			
AY108441.1			
GI:21211519			
KEYWORDS			
HTC.			
SOURCE			
Zea mays			
ORGANISM			
Zea mays			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.			
1 (bases 1 to 1514)			
REFERENCE			
AY108441			
LOCUS			
DEFINITION			
Zea mays P0083558 mRNA sequence.			
ACCESSION			
AY108441			
VERSION			
AY108441.1			
GI:21211519			
KEYWORDS			
HTC.			
SOURCE			
Zea mays			
ORGANISM			
Zea mays			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.			
1 (bases 1 to 1514)			
REFERENCE			
AY108441			
LOCUS			
DEFINITION			
Zea mays P0083558 mRNA sequence.			
ACCESSION			
AY108441			

<p>AUTHORS</p> <p>Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittatt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.</p> <p>TITLE</p> <p>Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes</p> <p>JOURNAL</p> <p>Unpublished (2002)</p> <p>REFERENCE</p> <p>2 (bases 1 to 1514)</p> <p>AUTHORS</p> <p>Coe, B.H.</p> <p>TITLE</p> <p>Direct Submission</p> <p>JOURNAL</p> <p>Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA</p> <p>COMMENT</p> <p>If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.</p>	
---	--

FEATURES
SOURCE

```

source
1..1514
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:634674"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont
Library"
/note="this sequence is part of a pro
cesses resulting from the applica
configs to seed Dupont configs; this
assembled by DuPont as part of a coll
overgo addressing of BACs in conjunct
Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.: 5,85e-30 Length: 1514
Score: 340.00 Matches: 97
Percent Similarity: 47.66% Conservative: 56
Best Local Similarity: 30.22% Mismatches: 96
Query Match: 23.79% Indels: 72
DB: 3 Gaps: 10
US-09-545-199F-4 (1-289) x AY108441 (1-1514)

```

[illegible]

Search completed: November 9, 2004, 21:31:27
Job time : 24324 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 9, 2004, 07:55:18 / Search time 639 Seconds
(without alignments)

2435.070 Million cell updates/sec

Title: US-09-545-199F-4

Perfect score: 1429

Sequence: 1 MAGAKEIRTKIASVKSTQKI.....ARQASITNELNEIVAGAAAI 289

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3611042 segs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cgn2_1/USPTO_pool_p/US09545199/runat_08112004_112527_21010/app_query.fasta_1.455
-DB=Published Applications_NA -GRMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl6sum62
-TRANS=human40.cdi -L-ST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09545199 @CGN 1.1 912 @runat 08112004 112527_21010
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THRAD=10 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1429	100.0	1972	11	US-09-809-665A-3
C 2	1173	82.1	1830121	14	US-10-329-960-1
C 3	1173	82.1	1830121	16	US-10-329-960-1
C 4	1173	82.1	1830121	18	US-10-158-865-1
5	1132.5	79.3	866	11	US-09-809-665A-132
6	1077.5	75.4	866	11	US-09-809-665A-166
7	1007	70.5	861	15	US-10-369-493-44931
8	1002	70.1	861	15	US-10-369-493-44931
9	991	69.3	864	9	US-09-741-669-144
10	991	69.3	864	15	US-10-369-493-47300
11	904.5	63.3	858	15	US-10-369-493-47300
12	891	62.4	867	15	US-10-369-493-32530
13	885.5	62.0	882	15	US-10-369-493-32530
14	885	61.9	94750	16	US-10-672-787-36
15	838	58.6	861	15	US-10-369-493-39109
16	838	58.6	861	15	US-10-369-493-39177
17	838	58.6	861	15	US-10-369-493-39177
18	812.5	56.3	876	15	US-10-369-493-39859
19	812.5	56.3	876	15	US-10-369-493-39859
20	809	56.6	861	15	US-10-369-493-30933
21	804	56.3	640681	9	US-10-369-493-33058
22	795	55.6	864	15	US-09-790-988-1
23	795	55.6	2731748	17	US-10-369-493-41252
C 24	787	55.1	861	15	US-10-297-465A-1
25	650.5	45.5	531	11	US-09-809-665A-97
26	585.5	41.0	843	15	US-10-369-493-32966
27	567.5	39.7	879	15	US-10-369-493-33589
28	536.5	37.5	858	15	US-10-369-493-32600
29	534.5	37.4	945	15	US-10-369-493-41141
30	530.5	37.1	864	15	US-10-369-493-43949
31	530.5	37.1	7750	9	US-10-369-493-47016
32	529	37.0	858	15	US-03-070-327A-216
33	524.5	36.7	4815	15	US-10-369-493-40239
34	523.5	36.6	870	15	US-10-310-630-1
35	518	35.7	851	9	US-09-974-300-100
36	509.5	35.7	945	15	US-10-369-493-43628
37	504.5	35.3	945	15	US-10-369-493-28299
38	504.5	35.3	2207	15	US-10-310-630-6
39	502	35.1	882	15	US-10-369-493-43236
40	501.5	35.1	888	17	US-10-474-776-529
41	501.5	35.1	10399	8	US-08-961-527-160
42	501.5	35.1	10399	16	US-10-158-844-160
43	489	34.2	684707	16	US-10-398-221-9
C 44	489	34.2	3011208	16	US-10-398-221-9
45	486	34.0	6309	8	US-08-781-986A-480

ALIGNMENTS

RESULT 1
US-09-809-665A-3
; Sequence 3, Application US/09809665A
; Publication NO. US20040110286A1
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28341/00435
; CURRENT APPLICATION NUMBER: US/09/809,665A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1972
; TYPE: DNA


```

/ LOCATION: (45593)..(45593)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (45732)..(45732)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (47036)..(47036)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (51334)..(51334)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (51602)..(51602)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (51786)..(51786)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (51805)..(51805)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (55369)..(55369)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (65309)..(65309)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (65313)..(65313)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (80024)..(80024)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (100091)..(100091)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (102696)..(102696)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (105121)..(105121)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (107248)..(107248)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (117136)..(117136)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (119750)..(119750)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (119924)..(119924)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (120038)..(120038)

```

```

/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (121344)..(121344)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (122167)..(122167)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (122336)..(122336)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (131340)..(131340)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (131360)..(131360)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (139910)..(139910)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (140398)..(140398)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (142750)..(142750)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145058)..(145058)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145171)..(145171)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145942)..(145942)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (147197)..(147197)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (150841)..(150841)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152500)..(152500)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152530)..(152530)

```

Alignment Scores:

Pred. No.:	7,62e-108	Length:	1830121
Score:	1173.00	Matches:	221
Percent Similarity:	91.00%	Conservative:	42
Best Local Similarity:	76.47%	Mismatches:	26
Query Match:	82.09%	Indels:	0
DB:	14	Gaps:	0

US-09-545-199f-4 (1-289) x US-10-329-960-1 (1-1830121)

QY

1 MetAlaGlyAlaLeysGluIleArgThrLysIleAlaSerVallySerThrGlnLysIle 20

Db 501934 ATGGCAGGTGCAAAAGAGATATAAAACCAAAATGCCAGTGTACAAAGTACACAAAAATC 501875
Qy 21 ThrlyAlaMetGluMetValAlaalaSerlyMetArglyThrGlnGluArgMetSer 40
Db 501874 ACTAAGGCAATGGAATGTTGGCAACCTCGGAATGGTAAACGAGGAGATCGTATGGCT 501815
Qy 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerlyAlaThr 60
Db 501814 GCATCTCGTCGATTCGTAATTCGAAACATATCCGTAACGTTATTAGTCAATGCTGAAGCAAGT 501755
Qy 61 IleGlyTyrIleHisProPheLeuValAspArgGluValIleValIleValGlyMetIleVal 80
Db 501754 ATCGGTTATAAACATCGTTCTTCTTAGTTGACGGCGAAGTAGGAAATACGGTATCTGGTT 501695
Qy 81 ValSerThrAspArgGlyLeuCysGlyIleAsnValAsnLeuPheIleThrValLeu 100
Db 501694 ATTTCAACAGATCGTGATGTGTGGTGGTAAATGTTAAATTTATTCACAAACACACTT 501635
Qy 101 AsnGluMetIleGlyLysGluLysAspValSerValGlnLeuSerIleLeuIleGlySer 120
Db 501634 AACCAATAAAAAATGGAAAGACAAATATTTCTACAGATTGGGCTTAATAGGTTCA 501575
Qy 121 LysSerIleAsnPhePheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db 501574 AAAGGATAGTTTTCGTTCTTCTTGGATTGAATATCAAGGTCAGCTTCTGGTTA 501515
Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db 501514 GCGATACGCCGCTCTAGAGAATTAATTGTGTGGCAATACATATGTTTGTGCTTAT 501455
Qy 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 501454 CGTAATGGTGAATGATGACGTTTATATGTCATACATAAATTTGTTAATACGATCG 501395
Qy 181 GlnLysProValLeuGluLysLeuIleProLeuGluLeuAspAsnAspGluLeuGly 200
Db 501394 CAAAGCCCTGTTGTACAAATAGTCTTACAGAACTAAAGACGATCTTTAAAT 501335
Qy 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db 501334 GAAAGACACAGCTGGGATTTATTTATGAGCAGACAAAGTACTATTAGATAGC 501275
Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
Db 501274 CTTTATGTTGTTTATGATGTCCTCAATTTATCAAGCGTTGTAGATATGATGCTTCA 501215
Qy 241 GluGlnAlaAlaArgMetValAlaMetIleValAlaThrAspAsnAlaGlyAsnLeuIle 260
Db 501214 GAAACAGCGCTCGAATGAGCAATGAAAGCAGCAGCGATATGCGGTAATTTAAT 501155
Qy 261 AsnGluLeuGlnLeuValTyrAsnValAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
Db 501154 AATGATCTGCGTTGGTGTACACAAAGCTCGTCAAGCAAGTATACAAATGAATGAAT 501095
Qy 281 GluIleValAlaGlyAlaAlaIle 289
Db 501094 GAAATCTAGCGGTGCGGCGGAT 501068
RESULT 3
US-10-329-670-1/c
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40810)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:

```

; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature

```

```

; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

```

Alignment Scores:

```

Pred. No.: 7.62e-108      Length: 1830121
Score: 1173.00           Matches: 221
Percent Similarity: 91.00%    Conservative: 42
Best Local Similarity: 76.47%  Mismatches: 26
Query Match: 82.09%          Indels: 0
DB: 16                     Gaps: 0

```

US-09-545-199F-4 (1-289) x US-10-329-670-1 (1-1830121)

```

Qy 1 MetAlaGlyAlaLysGluLeuArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
Db 501934 ATGGCAGGTGCAAAAGAGATAAAACCAAAATGCCAGTGTACAGTACACAAAAATC 501875
Qy 21 ThrIleAlaMetGluMetValAlaLaserLysMetArgLysThrGlnGluArgMetSer 40
Db 501874 ACTAAGGCAATGGAATGTGGTGGACCTCGAAAAATGCGTAAACGACGATCGTATGGCT 501815
Qy 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db 501814 GCATCTCGTCCGTATTCTGAAACTATCCGTAAACGTTATTAGTCATGTGCTTAAGCGCAGT 501755
Qy 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysValGlyMetIleVal 80
Db 501754 ATCGGTTATAACATCCGTTCTTAGTTCAGCCGCAAGTGAAGAAAAATCGTATCTTGGTT 501695
Qy 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db 501694 ATTCAACAGATCGTGGGATGTGGTGGTTAATGTTAATTTATTTCACAAACCACTT 501635
Qy 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120

```

```
Db 501634 AACCAATATAAAATTTGGAGACAAATATTTCTACAGATTTCGGCTTAATAGGTCA 501575
QY 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
Db 501574 AAAGGATAGTATTTTCGCTTCCTTGGATTAAATATCAAAAGTCAGCTTTCGGTTA 501515
QY 141 GlyAspThrProSerValGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db 501514 GCGGATACCCCGCTCTAGAGAAATTAATTTGGTGGCAATACAAATGTTTGATGCTTAT 501455
QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 501454 CGTAATGGTGAATGATGACGATTTATATTCATACAAATATTTGTTATACGATGTCG 501395
QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db 501394 CAABAGCTGTGTACACAAATAGTCTTTACAGAACTAAAGACGATCTTTAAAT 501335
QY 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db 501334 GAAGACACACAGCTGGGATATCTTTATGAGCCAGACAAAGTACTATTAGATAGC 501275
QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
Db 501274 CTTTATGTCGTTATTTAGATGTCCTCAATTTATCAGCGTTGTAGATAATGTAGCTTCA 501215
QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
Db 501214 GAACAAAGCGGTGCAATGATAGCAATGAACGACGACGATATGCGATTAATTAAT 501155
QY 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnLysIleThrAsnGluLeuAsn 280
Db 501154 AATGATCTGCGGTGGTGTACAAAGCTGTCRAAGCAAGTATCAAAATGAATGAAT 501095
QY 281 GluIleValAlaGlyAlaAlaIle 289
Db 501094 GAAATCGTAGCGGTGGCGGGCGGATT 501068

RESULT 4
US-10-158-965-1/c
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; FILE OF INVENTION: Thereof, and Uses Thereof
; CURRENT APPLICATION NUMBER: US/10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g

; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
```

```

FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (102596)..(102596)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119324)..(119324)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:

```

```

NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)

```

Alignment Scores:

```

Pred. No.: 7.62e-108 Length: 1830121
Score: 1173.00 Matches: 221
Percent Similarity: 91.00% Conservative: 42
Best Local Similarity: 76.47% Mismatches: 26
Query Match: 82.09% Indels: 0
DB: 18 Gaps: 0

```

US-09-545-199F-4 (1-289) x US-10-158-865-1 (1-1830121)

```

QY 1 MetalaglyalalysGluileArgThrLyleAlaserVallysserThrGlnLysile 20
Db 501934 ATGCAGGTGCAAAAGAGATAAAACCAAAATTCAGGTGTACAAAGTACACAAAATC 501875
QY 21 ThrLysalaMetGluMetValAlaAlaserLysMetArgLysThrGlnGluArgMetSer 40
Db 501874 ACTAAGGCATGGAATGGTGGCACTCGAAATCGTAAACCGCAGGATCGTATGGCT 501815
QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
Db 501814 GCATCTCGTCGTATCTGAAACTATCCGTAAAGTATTAGTCATGTGCTTAAGCAAGT 501755
QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysValGlyMetIleVal 80
Db 501754 ATCGGTATAAACATCCGTTCTTAGTGTAGCGCGAAGTGAAGAAATCGTATCTTGGT 501695
QY 81 ValSerThrAspArgGlyLeuGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
Db 501694 ATTTCAACAGATCGTGGGATGTGTGGTGGTAAATGTTAATTTATTCAAAACACACTT 501635
QY 101 AsnGluMetLysGluTyrLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
Db 501634 AACCAATATAAAATTTGGAAGACANAATATTTACAGATTTGGGCTTAATAGTTCA 501575
QY 121 LysSerIleAsnPheGlnSerLeuGlyIleLysileLeuThrGlnAspSerGlyIle 140
Db 501574 AAAGGATTAGTTTTTTCGGTTCTTGGATTATATCAAGGTACGCTTCTGGTTTA 501515
QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
Db 501514 GCGGATACGCCCGCTCTAGAAGAATTAATTCGTGTGCAAAATACAAATGTTTGAATGCTTAT 501455
QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
Db 501454 CGTAATGGTGAATTTGATGCAATTTATTCATPACAAATAAATTTGTTAATACGATGCG 501395
QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
Db 501394 CAAAAGCCTGTTGTACAAACATTAGTTCCTTTACCAGATCTTAAGACGATCATTTAAAT 501335

```

QY 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAen 220
 Db 501334 GAAAGACACAGACTTGGGATTATCTTATGACCCAGAACCAAAAGTAGTATTAGATAGC 501275
 QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
 Db 501274 CTTTATTGTTATTAGAGTCCCAAAATTATCAAGCGGTGTAGATAATGTAGCTTCA 501215
 QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAenAlaGlyAsnLeuIle 260
 Db 501214 GAACAGCGGCTCGAATGGTAGCAATGAAAGCAGCAACGATAATGTCAGGTAATTAAAT 501155
 QY 261 AsnGluLeuGlnValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAen 280
 Db 501154 AATGATCTCGGTTGGTGTACACAAAGCTCGTCAAGCAAGTATCACAAATGAATTGAAT 501095
 QY 281 GluIleValAlaGlyAlaAlaLalle 289
 Db 501094 GAAATCGTAGCGGTCGCGCGCGATT 501068
 RESULT 5
 ; Sequence 132, Application US/09809665A
 ; Publication No. US20040110268A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lowery E., David, et al.
 ; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
 ; FILE REFERENCE: 28341/00435
 ; CURRENT APPLICATION NUMBER: US/09/809,665A
 ; PRIORITY FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 60/153,453
 ; PRIOR FILING DATE: 1999-09-10
 ; PRIOR APPLICATION NUMBER: 60/128,689
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 09/545,199
 ; NUMBER OF SEQ ID NOS: 197
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 132
 ; LENGTH: 867
 ; TYPE: DNA
 ; ORGANISM: Actinobacillus pleuropneumoniae
 ; FEATURE:
 ; OTHER INFORMATION: atpG
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(864)
 ; US-09-809-665A-132
 Alignment Scores:
 Pred. No.: 2,66e-108 Length: 867
 Score: 1132.50 Matches: 215
 Percent Similarity: 89.97% Conservative: 45
 Best Local Similarity: 74.39% Mismatches: 28
 Query Match: 79.25% Indels: 1
 DB: 11 Gaps: 1
 US-09-545-199f-4 (1-289) x US-09-809-665A-132 (1-867)
 QY 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
 Db 1 ATGCGAGGTGCGAAGAGATAGAACCAAAATGCAAGTGTGAAATACTCAAAAATC 60
 QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnLysArgMetSer 40
 Db 61 ACCAAAGCAATGAAATGGTTGTCTACCTCTAAATGCGTAAACCGCAGAGCGTATGGCT 120
 QY 41 SerSerArgProTyrSerGlnThrIleArgAsnValIleSerHisValSerLysAlaThr 60
 Db 121 GCCAGTCGTCCTTATTCGGAACAATCCGTAAAGGTATAGCCATATTGCGAAAGGAGC 180
 QY 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
 Db 181 ATTGGTTAAGCACCCTGTTTAACTCAAGCTGATTAATAAAGTAGGCTATCTTGTC 240

QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAenValAsnLeuPheLysThrValLeu 100
 Db 241 GTTTCGACCGATCGCGTGTATTCGCGTGCCTTAATATCAATTTATTCAAAGCCACTTGG 300
 QY 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
 Db 301 AATGAAATTTAAACCTGGAAGATTAAGACCTTGTGTGACCTTGGTTAGTAGGCGTCG 360
 QY 121 LysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
 Db 361 AAAGCGTAAGCTTTTACCAAAATCTAGGCTTAACGTGAGATCTCAAGTAACGGGATTA 420
 QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
 Db 421 GCGGATAATCCGGAATGGAACGTATCGTGGCGCAGTTAATGAAATGATTATATGCGTTC 480
 QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
 Db 481 CGAAACGGAGAGTGGATGCGGTTTACGTGCTTACACCGTTTTGAAATACGATGTCA 540
 QY 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
 Db 541 CAAAAACCTGTTATCGCACAGTACTTCCGTTACTTAACTAACTAGATGACGATGAATTA--- 597
 QY 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAen 220
 Db 598 GATACGAAAGGTTTCATCGGATTATATTTATGAACCGAATCCACAGTTTATTTGATAGT 657
 QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
 Db 658 TTACTTGTCTGTTATTAGAACTCAGGTATACCAAGCAGTTGTAGATAACCTACTTCT 717
 QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
 Db 718 GAACAAGCCGCTCGAATGCTAGCGATGAAGCGGCAACAGATAATGCGGTTACATTAATC 777
 QY 261 AsnGluLeuGlnValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAen 280
 Db 778 GATGAATTAATTAATGATGTTATTAACAAAGCTCCCAAGCAAGCATTACAAATGAATTAAC 837
 QY 281 GluIleValAlaGlyAlaAlaLalle 289
 Db 838 GAAATGTTGCGGTCGCGCGCGCAATTT 864
 RESULT 6
 ; Sequence 166, Application US/09809665A
 ; Publication No. US20040110268A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lowery E., David, et al.
 ; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
 ; FILE REFERENCE: 28341/00435
 ; CURRENT APPLICATION NUMBER: US/09/809,665A
 ; PRIORITY FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 60/153,453
 ; PRIOR FILING DATE: 1999-09-10
 ; PRIOR APPLICATION NUMBER: 60/128,689
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 09/545,199
 ; NUMBER OF SEQ ID NOS: 197
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 166
 ; LENGTH: 866
 ; TYPE: DNA
 ; ORGANISM: Pasteurella (Mannheimia) haemolytica
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(864)
 ; OTHER INFORMATION: atpG
 ; US-09-809-665A-166

Db 598 ---AAGAGAAGTCCTGGGATATCTGTATGAACCTGATCCTAAGCGGTGTGGATACC 654
 Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluLeuLeuAlaSer 240
 Db 655 CTGCTGCGCGCTATATAGAAATCACAGTTTATCAGGGCGTGTGTAAGAAACCTGGCTAGT 714
 Qy 241 GluGlnAlaAlaArgMetValAlaMetLysAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuLeu 260
 Db 715 GAACAGCGCCGACCAATGGTAGCGATGAAGAGCGCGACCGGTAACGGTGGGACCTGATC 774
 Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 Db 775 AAGAGGTTCAGTGTGGTTTACAAAGGCTGTCAGGCGCAGCATCCTCAGGAATCACC 834
 Qy 281 GluLeuValAlaGlyAlaAlaLeu 289
 Db 835 GAAATCGTTTCGGGTGCTTCGGGTT 861
 RESULT 8
 US-10-369-493-24123
 ; Sequence 24123, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 473/4
 ; SEQ ID NO 24123
 ; LENGTH: 861
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus nematophilus
 US-10-369-493-24123

Alignment Scores:
 Pred. No.: 1.03e-94 Length: 861
 Score: 1002.00 Matches: 194
 Percent Similarity: 83.74% Conservative: 48
 Best Local Similarity: 67.13% Mismatches: 45
 Query Match: 70.12% Indels: 2
 DB: 15 Gaps: 1

US-09-545-199F-4 (1-289) x US-10-369-493-24123 (1-861)
 Qy 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
 Db 1 ATGCCCGCGCGCAAAAGAGATACGTACCAATCGCAGTGTGCAGAACACGCGAAGATC 60
 Qy 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
 Db 61 ACTAATACCGATGAAATGGTGGCGCGGTCCAAATCGTAAACGCGACGCGCATGGCG 120
 Qy 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
 Db 121 GCCAGCGCTCTTATGACATACCATTCGACCGTATGGACACTTGGCTGGCTAGTAAAT 180
 Qy 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysValGlyMetIleVal 80
 Db 181 CTGGAGTACAAACACCCATACCTTGAAGAGCGGTGACGTTAAACGGGTGGGTACGTGTT 240
 Qy 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrValLeu 100
 Db 241 GTTCTTACTACCGTGTGTTATGTGCGGTTTGACATTAATCTGTTCAAAAAATTCGTG 300
 Qy 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerIleGlySer 120

Db 301 ATAGAATGAAGACTGGTCTGATATAAAGAGTCCAGGTGTGATTTGGCACTATTGGATCA 360
 Qy 121 LysSerIleAsnPhePheGlnSerLeuGlyLysIleLysIleLeuThrGlnAspSerGlyIle 140
 Db 361 AAGCGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420
 Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
 Db 421 GGAGATAACCCATCATCTGCCAATTGATCGGCCGACGTCATGTCATGATCAAGCATAT 480
 Qy 161 LysLysGlyGluValAspValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
 Db 481 GACGAGGGCGTCTGGATAAACTGTATATAGTACAAACAAAGTTCCACAATACATGCTCT 540
 Qy 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
 Db 541 CAGGTTCCGACAACTCAGTCTATTCCTCTGCGCGGAGACGATGAACACTG--- 597
 Qy 201 GluArgLysGlnValTyrAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
 Db 598 ---AAGAGAAGTCTCTGGGATATCTGTATGAACCATCTTAAGGCGCTGTGGATACC 654
 Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluLeuLeuAlaSer 240
 Db 655 TTGCTGCGTCTATATAGAAATCGCAAGTTTATCAGGGCGTGTGTAAGAAACCTGGCTAGT 714
 Qy 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuLeu 260
 Db 715 GAACAGCGCGCAGCAATGGTAGCGATGAAGCGCGACTGATTAACGTCGCAACCTGATC 774
 Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 Db 775 AAGAGTTGCGTGTGGTTTACAAAGCTCGTCAGGCGCAGCATTAATCAGGAATCACC 834
 Qy 281 GluLeuValAlaGlyAlaAlaLeu 289
 Db 835 GAAATCGTTTCGGGTGCGCTGGGTT 861

RESULT 9

US-09-741-669-144
 ; Sequence 144, Application US/09741669
 ; Patent No. US20020022718A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Forsyth, R. Allyn
 ; APPLICANT: Olesen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; TITLE OF INVENTION: Genes identified as required for
 ; TITLE OF INVENTION: proliferation of E. coli
 ; FILE REFERENCE: ELITRA.009A
 ; CURRENT APPLICATION NUMBER: US/09/741,669
 ; CURRENT FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: US 60/173005
 ; PRIOR FILING DATE: 1999-12-23
 ; NUMBER OF SEQ ID NOS: 481
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 144 -
 ; LENGTH: 864
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(864)
 US-09-741-669-144

Alignment Scores:
 Pred. No.: 1.44e-93 Length: 864
 Score: 991.00 Matches: 191
 Percent Similarity: 82.35% Conservative: 47
 Best Local Similarity: 66.09% Mismatches: 49
 Query Match: 69.35% Indels: 2
 DB: 5 Gaps: 1


```

US-09-545-199f-4 (1-289) x US-09-741-669-144 (1-864)
QY 1 MetAlaGlyAlaLysGluLeuArgThrLysLysLeuAlaSerValLysSerThrGlnLysIle 20
Db 1 ATGGCGCGCGCAAAAGAGATAGTAGTAAGATCGCAGCGCTCCAGACACCGCAAAAGATC 60
QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
Db 61 ACTAAAGCGATGGAGATGGTGGCGCTCCAAATGCGTAATCGCAGGATCGCATGGCG 120
QY 41 SerSerArgProTyrSerGluThrIleArgAsnValLysSerHisValSerLysAlaThr 60
Db 121 GCCAGCGCTCTTATGAGAACCATCGCAAGATGATGGTACCTTGCACCGGTAAT 180
QY 61 IleGlyTyrLysHisProPheLeuValAspValSerValGlnLeuSerLysIleGlySer 120
Db 301 GCGAAATAGACACCTTACCTGGAAGACCGGACGCTTAAACGCGTGGCTACCTGGTG 240
QY 81 ValSerThrAspArgGlyLeuGlyLysGluValAspValSerValGlnLeuSerLysIle 160
Db 361 AAAGCGCTGCTTCTTCAACTCCGTCGCGGCGCATGTTGTTGCCCGCATGTTTCAAAAACCTGCTG 420
QY 141 GlyAspThrProSerValGluGlnLeuLysLeuGlySerValAsnSerMetIleAspAlaTyr 160
Db 421 GGGGATTAACCTTCCCTGTCGGAACCTGATCGGTCCGGTAAAGTATGTTGCAGCCATC 240
QY 161 LysSerIleAsnPheGlnSerLeuGlyLysGluValAspValSerValGlnLeuSerLysIle 140
Db 361 AAAGCGCTGCTTCTTCAACTCCGTCGCGGCGCATGTTGTTGCCCGCATGTTTCAAAAACCTGCTG 420
QY 141 GlyAspThrProSerValGluGlnLeuLysLeuGlySerValAsnSerMetIleAspAlaTyr 160
Db 421 GGGGATTAACCTTCCCTGTCGGAACCTGATCGGTCCGGTAAAGTATGTTGCAGCCATC 240
QY 161 LysSerIleAsnPheGlnSerLeuGlyLysGluValAspValSerValGlnLeuSerLysIle 140
Db 361 AAAGCGCTGCTTCTTCAACTCCGTCGCGGCGCATGTTGTTGCCCGCATGTTTCAAAAACCTGCTG 420
QY 181 GlnLysProValLeuGluValAspValSerValGlnLeuSerLysIleGlySer 120
Db 541 CAGGTTCCGACCATCAGCCAGCTGCTGCGGTACCGGCATCAGATGATGATGATCTG--- 597
QY 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db 598 ---AAACATAATCTGGGATTTACCTGACGAACCGCATCCGAAAGCGGCTTCTGGATACC 654
QY 221 LeuLeuValArgTyrLysLeuGluSerGlnValTyrGlnAlaLysValLeuLeuAspAsn 240
Db 655 CTGCTCGCTGCTATGTCGAATCTCAGGTTTATCAGGCGCTGTTTGAACACCTGGCCAGC 714
QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaLysAlaThrAspAsnAlaGlyAsnLeuIle 260

```

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

```

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 47300
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-47300

```

```

Alignment Scores:
Pred. No.: 1,44e-93 Length: 864
Score: 991.00 Matches: 191
Percent Similarity: 82.35% Conservative: 47
Best Local Similarity: 66.09% Mismatches: 49
Query Match: 69.35% Indels: 2
DB: 15 Gaps: 1

```

```

US-09-545-199f-4 (1-289) x US-10-369-493-47300 (1-864)
QY 1 MetAlaGlyAlaLysGluLeuArgThrLysLysLeuAlaSerValLysSerThrGlnLysIle 20
Db 1 ATGGCGCGCGCAAAAGAGATAGTAGTAAGATCGCAGCGCTCCAGACACCGCAAAAGATC 60
QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
Db 61 ACTAAAGCGATGGAGATGGTGGCGCTCCAAATGCGTAATCGCAGGATCGCATGGCG 120
QY 41 SerSerArgProTyrSerGluThrIleArgAsnValLysSerHisValSerLysAlaThr 60
Db 121 GCCAGCGCTCTTATGAGAACCATCGCAAGATGATGGTACCTTGCACCGGTAAT 180
QY 61 IleGlyTyrLysHisProPheLeuValAspValSerValGlnLeuSerLysIleVal 80
Db 181 CTGGAATATAGACACCTTACCTGGAAGACCGGACGCTTAAACGCGTGGCTACCTGGTG 240
QY 81 ValSerThrAspArgGlyLeuGlyLysGluValAspValSerValGlnLeuSerLysIle 100
Db 241 GTGTCGACCGCGCTGTTGTCGGTGGTGGTGAACATTAACCTGTTCAAAAACCTGCTG 300
QY 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLysIleGlySer 120
Db 301 GCGAAATAGACACCTTACCTGGAAGACCGGACGCTTCAATCGACCTCGCAATGATCGCTCG 360
QY 121 LysSerIleAsnPheGlnSerLeuGlyLysGluValAspValSerValGlnLeuSerLysIle 140
Db 361 AAAGCGCTGCTTCTTCAACTCCGTCGCGGCGCATGTTGTTGCCCGCATGTTTCAAAAACCTGCTG 420
QY 141 GlyAspThrProSerValGluGlnLeuLysLeuGlySerValAsnSerMetIleAspAlaTyr 160
Db 421 GGGGATTAACCTTCCCTGTCGGAACCTGATCGGTCCGGTAAAGTATGTTGCAGCCATC 480
QY 161 LysSerIleAsnPheGlnSerLeuGlyLysGluValAspValSerValGlnLeuSerLysIle 180
Db 481 GACGAAGCGCTGTCGCAAGCTTTACATTGTCAGCAACAAATTTATTAAACACCTGCTCT 540
QY 181 GlnLysProValLeuGluValAspValSerValGlnLeuSerLysIleGlySer 200
Db 541 CAGGTTCCGACCATCAGCCAGCTGCTGCGGTACCGGCATCAGATGATGATGATCTG--- 597
QY 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
Db 598 ---AAACATAATCTGGGATTTACCTGACGAACCGCATCCGAAAGCGGCTTCTGGATACC 654
QY 221 LeuLeuValArgTyrLysLeuGluSerGlnValTyrGlnAlaLysValLeuLeuAspAsn 240
Db 655 CTGCTCGCTGCTATGTCGAATCTCAGGTTTATCAGGCGCTGTTTGAACACCTGGCCAGC 714
QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaLysAlaThrAspAsnAlaGlyAsnLeuIle 260

```

RESULT 10

US-10-369-493-47300

Sequence 47300, Application US/10369493

Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

Db 715 GAGCAGCCGCCGCTATGGTGGGATGAAAGCCGCGACCAATGGCGGAGCTGATT 774
 Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 Db 775 AAAGAGCTGCGAGTTGGTATACAAAGAGCTGTCAGGCAGCATTTACTCAGGAATCACC 834
 Qy 281 GluIleValAlaGlyAlaAlaAla 289
 Db 835 GAGATCGTCTCGGGGCGCGCGGTT 861
 RESULT 11
 US-10-369-493-37532
 ; Sequence 37532, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 37532
 ; LENGTH: 858
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas fluorescens
 US-10-369-493-37532
 Alignment Scores:
 Pred. No.: 1.45e-84 Length: 858
 Score: 904.50 Matches: 177
 Percent Similarity: 78.20% Conservative: 49
 Best Local Similarity: 61.25% Mismatches: 60
 Query Match: 63.30% Indels: 3
 DB: 15 Gaps: 1
 US-09-545-199F-4 (1-289) x US-10-369-493-37532 (1-858)
 Qy 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
 Db 1 ATGCAGCGCCAAAGAGATTCGAGTAAGATTGCGAGCATCAAAGCAGCAGCAAAATT 60
 Qy 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnLysArgMetSer 40
 Db 61 ACCAGCGCCATGAAAGAGTGGCGGTACGAAATGCGCAGGCAAAATGCGCATGCGT 120
 Qy 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
 Db 121 GCTAGCGCGTCTTATCGGAGCGGTATCCGCCAGGTAATGGGCATCTCGCCAAACGCCAAC 180
 Qy 61 IleGlyTyrLysHisProPheLeuValAspArgGluValLysLysValGlyMetIleVal 80
 Db 181 CCGGAATACCGCCACCGCTTCATGATCGATCCGGAAGTAAAGCGCGGTGTTATGTCGTA 240
 Qy 81 ValSerThrAspArgGlyLeuGlyLysGlyLeuAsnValAsnLeuPheLysThrValLeu 100
 Db 241 GTGAGCAGTGACCGTGGTGTGTGCGCGCGCTTGACACCAACCTGTTCAAGGCCCTGGTC 300
 Qy 101 AsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGlySer 120
 Db 301 AAGGACATGGCGGTAAACCGCGAAACCGCGTCGAGATCGATCTGTGTCGTGGTAGC 360
 Qy 121 LysSerIleAsnPhaPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
 Db 361 AAGGTGGCGCGCTTTTCCGAACTTCGCGGTAACTGCGTTCGAGCATCACGCCACCTG 420
 Qy 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160

Db 421 GGTGAAGAGCCGCTGATCAATCATCTGATCGCAGCGTCAAGGTGATGCTGGATGCCTAT 480
 Qy 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
 Db 481 CTGGACGCGCGTATTGACCGCTGTCTGGTGTATCCCAACAAGTTTCATCAACACCATGACG 540
 Qy 181 GlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
 Db 541 CAACAGCCTACCGTGGAGCAGTTGATTCATCCATGGTGGCGACCCCGGATCAAGAACTC--- 597
 Qy 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220
 Db 598 -----AAGCACCACTGGGACTACCTCTACGAAACCGACGCCAAGAGTGTGTTGACGC 651
 Qy 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGluAlaAlaValGluLeuAsnLeuAsp 240
 Db 652 TTGATGGTGGCTTACGTGGAGTCCGAGGTGTACAGCGGTGTCGAGAACCAACGCGGT 711
 Qy 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuIle 260
 Db 712 GAACAGCTCGCGGATGATCGGATGAAGAACCTTACCGACAACGCCGTGATTGATC 771
 Qy 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 Db 772 AGCGATTGCGAGCTGATCTACAAAGCGCGTACCGACAACGCCGTGATTGATC 831
 Qy 281 GluIleValAlaGlyAlaAlaAla 289
 Db 832 GAAATCGTGGCGCGCTGCGCGGTT 858

RESULT 12

US-10-369-493-32530
 ; Sequence 32530, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Chen, Xianfeng
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 FILE REFERENCE: 38-10(52052)B
 CURRENT APPLICATION NUMBER: US/10/369,493
 CURRENT FILING DATE: 2003-02-28
 PRIOR APPLICATION NUMBER: US 60/360,039
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 47374
 SEQ ID NO 32530
 LENGTH: 867
 TYPE: DNA
 ORGANISM: Ralstonia metallidurans
 US-10-369-493-32530

Alignment Scores:
 Pred. No.: 3.75e-83 Length: 867
 Score: 891.00 Matches: 171
 Percent Similarity: 78.89% Conservative: 57
 Best Local Similarity: 59.17% Mismatches: 59
 Query Match: 62.35% Indels: 2
 DB: 15 Gaps: 1

US-09-545-199F-4 (1-289) x US-10-369-493-32530 (1-867)

Qy 1 MetAlaGlyAlaLysGluIleArgThrLysIleAlaSerValLysSerThrGlnLysIle 20
 Db 1 ATGCCTGGACGAAAGAGATTCGAACCAAGATCAAGAGCGTGCACAAACACGCGCAAGATC 60
 Qy 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnLysArgMetSer 40
 Db 61 ACCAAGCGCGATGAGATGTCGCGCGATCCAGATCGCGCAGGCGCAGGACGATGCGC 120

QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerIysAlaThr 60
 DB 121 AACGCCGCCCTACGCGGAGAAAGTCGGAATATCCGCGCACCTGGCCCTGCGCAAC 180
 QY 61 IleGlyTyrIysHisProPheLeuValAspArgLysValIysValGlyMetIleVal 80
 DB 181 CCCAGATTCAAGCATCGCTTCATGGTGGCAGCGGAGCTCAAGCGTCCGCGCATGATCGTG 240
 QY 81 ValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheIysThrValLeu 100
 DB 241 GTGACGACCCACCAAGGTCGTGCGGTGGCTGACACCAACGTCGTGCGCGGTGACC 300
 QY 101 AsnGluMetIysGluTrpIysAspValSerValIleLeuSerIleGlySer 120
 DB 301 AACCAACTGAAGACCTGCAGCGGCCAGCGGTGAACGTGCAAGCGACCGCATCGGTACC 360
 QY 121 LysSerIleAsnPhePheGlnSerLeuGlyIleIleLeuThrGlnAspSerGlyIle 140
 DB 361 AAGGGCATGAGTTCCTGGCCCATCGCGCCAGAGTGGTCTCCGATGTGGTGCAGGTG 420
 QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaTyr 160
 DB 421 GGTGACACCCCGCATCTGGAAGCTGATCGCGCGGATCAAGTTTCACCTGCGCCCTAT 480
 QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
 DB 481 ACCAACGGGAAGTCTGACCGGTGACTTGGCATATACCAAGTTTCATCAACACGATGAAG 540
 QY 181 GlnLysProValLeuGluLysLeuIleProLeu-----ProGluLeuAspAsnAspGlu 198
 DB 541 CAGGAGCCGATGTCGAGCACTGTCGCGCTGGCAGCGGCAAGTTCAGCCAGCCGACCGAA 600
 QY 199 LeuGlyGluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeu 218
 DB 601 GATGAAGCGCGCTACTCGTGGGATATACATCTACGAGCCGCGCAGCGCCAGCGTTGTG 660
 QY 219 AspAsnLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeu 238
 DB 661 GAAGAGTGTCTGTCGCTAGCTGCAAGCGTGTGTACAGCGCGTGGCGGAGAACATG 720
 QY 239 AlaSerGluGlnAlaAlaArgMetValAlaMetIysAlaAlaThrAspAsnAlaGlyAsn 258
 DB 721 CGCTCGGAGCAATCCGCGCTATGTCGCTGATGAGCGTGTGTACAGCGCGTGGCGGAGAACATG 780
 QY 259 LeuIleAsnGluLeuGlnValTyrAsnLysAlaAlaArgGlnAlaSerIleThrAsnGlu 278
 DB 781 GTGATCGCGCACTGCACTGCTCTACAAAGACCCGTCAGGACGCGATACGAAGAA 840
 QY 279 LeuAsnGluIleValAlaGlyAlaAla 287
 DB 841 CTGTGGAAATCGTCAGCGGTGCGCT 867

RESULT 13

US-10-369-493-43286
 ; Sequence 43286, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 43286
 ; LENGTH: 862
 ; TYPE: DNA

; ORGANISM: Nitrosomonas europaea
 US-10-369-493-43286

Alignment Scores:

Pred. No.: 1,44e-82 Length: 882
 Score: 885.50 Matches: 165
 Percent Similarity: 78.91% Conservative: 67
 Best Local Similarity: 56.12% Mismatches: 57
 Query Match: 61.97% Indels: 5
 DB: 15 Gaps: 2

US-09-545-199F-4 (1-289) x US-10-369-493-43286 (1-882)

QY 1 MetAlaGlyAlaLysGluIleArgThrIysIleAlaSerValLysSerThrGlnLysIle 20
 DB 1 ATGCCCAGCAGCAGAGATAGCGCAATAAATCAAAACGGTTAAAAATACCACAAAGATT 60
 QY 21 ThrLysAlaMetGluMetValAlaAlaSerLysMetArgLysThrGlnGluArgMetSer 40
 DB 61 ACACGTGCCCATGGAATGGTTGCCCTTCCAAATAGGAAGGCCAGGATCGTATGAAG 120
 QY 41 SerSerArgProTyrSerGluThrIleArgAsnValIleSerHisValSerLysAlaThr 60
 DB 121 AAAGCGCTCCTTATGGCGAGAGATAGCAACGTGGCTGCACATATGAGTAATGCCAGT 180
 QY 61 IleGlyTyrLysHisProPheLeuValAspArgLys---ValLysLysValGlyMetIle 79
 DB 181 GTTGTAGTATCCCATCCCTTTTCTGATAAGTCTGACTCTGTATAAAGAGTGGGATCATT 240
 QY 80 ValValSerThrAspArgGlyLeuCysGlyGlyLeuAsnValAsnLeuPheLysThrVal 99
 DB 241 GTTGTAACTTCAGATAAAGTTTGTGTGGCGGTGAAACACGATGTGCTGCATAGCA 300
 QY 100 LeuAsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuIleGly 119
 DB 301 CTCAATGAATTCGAAGCTGGAAACCGAAGGTAATCATGTAGTCTGCTGCATAGGT 360
 QY 120 SerLysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGly 139
 DB 361 AACAGGGTTGGGTTTCATGACCGCTTGGTACTCAGGTAATCTCACAAAGTAACCGGG 420
 QY 140 IleGlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAla 159
 DB 421 TTGGTGTATGCACCTAATATGGAAGGTTGATTTGGTGCAGTCAAGTAGTCTTGGATGCA 480
 QY 160 TyrLysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMet 179
 DB 481 TATACAGAGCGCAGTTGGATAGGTTATATTTCTATAACCGCTTTATTAATACAATG 540
 QY 180 SerGlnLysProValLeuGluLysLeuIleProLeuProGlu-----LeuAsp 195
 DB 541 AAGCAGATGCGGTTATGGAGCAGCTTCTCCCTTGATGATGATCGTATCAGTAGTGAG 600
 QY 196 AsnAspGluLeuGlyGluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLys 215
 DB 601 GATGGGAAAGCCAGACCTACCGCTGCACCGTGGATTATATTTATGAACCTGAGCAAAA 660
 QY 216 ValLeuLeuAspAsnLeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaVal 235
 DB 661 CCGGTTATCGCAGATATCATGTCAGGTATATAGAGGCGCTCGTATACCGGAGTTGCT 720
 QY 236 GluAsnLeuAlaSerGluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsn 255
 DB 721 GAGACATGCTTCAGAACTTCGCAAGGATGGTTGCAATGAAGAGCGGCATCTGATAAT 780
 QY 256 AlaGlyAsnLeuIleAsnGluLeuGlnValTyrAsnLysAlaArgGlnAlaSerIle 275
 DB 781 GCAGGTAACTGATTGATGAGCTGACGCTCAATTATACCAAGTCTCTGCGAGGCCATT 840
 QY 276 ThrAsnGluLeuAsnGluIleValAlaGlyAlaAlaIle 289
 DB 841 ACCAAGGAATTTGCAGAAATCGTGGCGGTGCTGCTGCTGTT 882

RESULT 14

US-10-672-787-38
 ; Sequence 38, Application US/10672787
 ; Publication No. US20040067554A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LAGACE, Robert, E.
 ; APPLICANT: PATTERSON, Chandra
 ; APPLICANT: BERG, Kim, L.
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
 ; FILE REFERENCE: EPITRA.025C1
 ; CURRENT APPLICATION NUMBER: US/10/672,787
 ; PRIOR FILING DATE: 2003-09-26
 ; PRIOR APPLICATION NUMBER: 09/596,002
 ; PRIOR FILING DATE: 2000-06-16
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 38
 ; LENGTH: 94750
 ; TYPE: DNA
 ; ORGANISM: Moraxella catarrhalis
 US-10-672-787-38

Alignment Scores:

Pred. No.: 1,16e-79 Length: 94750
 Score: 885.00 Matches: 164
 Percent Similarity: 78.55% Conservative: 63
 Best Local Similarity: 56.75% Mismatches: 62
 Query Match: 61.93% Indels: 0
 DB: 16 Gaps: 0

US-09-545-199F-4 (1-289) x US-10-672-787-38 (1-94750)

QY 1 MetAlaGlyAlaLeuGluThrLysSerValLysSerThrGlnLysLeu 20
 Db 74372 ATGGCAAGCTTAAAGAAATTCGTCTAAGTACCGACGATTAAGATCGCAAGAT 74431
 QY 21 ThrLysAlaMetGluMetValAlaLysSerLysMetArgLysThrGlnLysMetSer 40
 Db 74432 ACGGTGCAATGCGATGCGTGGCTGCCAGTAAGATGCGTCCCAAGAGAGATGGAA 74491
 QY 41 SerSerArgProTyrSerGluThrIleArgAsnValLysSerHisValSerLysAlaThr 60
 Db 74492 TTGGTCTGCTCTTATTCAGATGGTATACCGCGTGTCAATTCACATTTGGTGCAGCGCAA 74551
 QY 61 IleGlyTyrLysHisProPheLeuValAspValSerValGlnLysSerLysLeuGlySer 80
 Db 74552 TCAGATTATAGCATCCCTATATGATTAACCGTCCGTCATCGCTTGGTTTGGTT 74611
 QY 81 ValSerThrAspArgGlyLeuGlyGlyLeuAsnValAsnValAsnLysThrValLeu 100
 Db 74612 GTAACCTCTGACCGTGGTTGGCAGGTGGTTTGACATCAATCTATTCAAAACACTGTGA 74671
 QY 101 AsnGluMetLysGluThrLysGluLysAspValSerValGlnLysSerLysLeuGlySer 120
 Db 74672 AAAACAGCTCAAAAGCTATCAAGACGACGTGAGTGAATTAAGTTCAGTAAATAGTGCA 74731
 QY 121 LysSerIleAsnPhePheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGlyIle 140
 Db 74732 AAAGCGTGGCGCTTTTAAAGACTTTGGCGTAGAGTAACCTCAGCTGTCACTACTAT 74791
 QY 141 GlyAspThrProSerValGluGlnLeuIleGlySerValAsnSerMetIleAspAlaThr 160
 Db 74792 GCGGTAACCTCCATTCAGCAAAATTAATACACCATTCACCATCTGTGATGATAT 74851
 QY 161 LysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMetSer 180
 Db 74852 CTCAATGGAAATAGATCGCATTTACTTGGTGTATACCATTCATCAATGCTATGGCT 74911
 QY 181 GlnLysProValLeuGlnLysLeuIleProLeuProGluLeuAspAsnAspGluLeuGly 200
 Db 74912 CAAAAACCGGTGTTGACAGATTGTACCATTTGGCAGAGAGTGTAGTTGATGACACTGAG 74971
 QY 201 GluArgLysGlnValTrpAspTyrIleTyrGluProAspAlaLysValLeuLeuAspAsn 220

Db 74972 CTTCAGGCACACAGTTGGGATTATATTTATGAGCCGACACCAAAACCGCTCATTGATAGC 75031
 QY 221 LeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGluAsnLeuAlaSer 240
 Db 75032 CTACTATTGGTTACATTGAATCTGTAGCTACCAATCCGTCGAGAGAAATTGCGCTCT 75091
 QY 241 GluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAlaGlyAsnLeuLeu 260
 Db 75092 GAACAGTCAGCCCGTATGTCGGGATGAACAGCGCAGCATATGCTGGTAACCTTAAT 75151
 QY 261 AsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThrAsnGluLeuAsn 280
 Db 75152 AAAGATTACAAATGGTTTATAAAGCTCAGACAGCGGCGATTACACGAGAAATCTCG 75211
 QY 281 GluIleValAlaGlyAlaAlaAlaIle 289
 Db 75212 GAAATCGTTGGTGGTGGCTGCCGCTGTT 75238

RESULT 15

US-10-369-493-39109
 ; Sequence 39109, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 39109
 ; LENGTH: 861
 ; TYPE: DNA
 ; ORGANISM: Xanthomonas campestris
 US-10-369-493-39109

Alignment Scores:
 Pred. No.: 1.23e-77 Length: 861
 Score: 838.00 Matches: 158
 Percent Similarity: 77.47% Conservative: 69
 Best Local Similarity: 53.92% Mismatches: 56
 Query Match: 58.64% Indels: 10
 DB: 15 Gaps: 3

US-09-545-199F-4 (1-289) x US-10-369-493-39109 (1-861)

QY 1 MetAlaGlyAlaLeuGluThrLysSerValLysSerThrGlnLysLeu 20
 Db 1 ATGGCAGCGCGGACGGAATCAACCAAGATCAAGCGTGCAGAACACCCGCAAGGTG 60
 QY 21 ThrLysAlaMetGluMetValAlaLysSerLysMetArgLysThrGlnLysMetSer 40
 Db 61 ACGCGCGCGCTCGAAATGCTCGGCGCTCCCAAGATCCGCAAGCGCAGGACCGCATGAAG 120
 QY 41 SerSerArgProTyrSerGluThrIleArgAsnValLysSerHisValSerLysAlaThr 60
 Db 121 ACCTCGCTCCCTACGCGCAGGCGCATGAACAGGTGTATCGGCGCATCTGGCGCAGGCGCAGC 180
 QY 61 IleGlyTyrLysHisProPheLeuValAspArgLysValLysValGlyMetIle 79
 Db 181 ACCGACTACCAAGCATCGCTTCTCGTGGAGCGTGAAGCAGTCAAGCGGTTCGCTACATC 240
 QY 80 ValValSerThrAspArgGlyLeuGlyGlyLeuAsnValAsnValAsnLysThrVal 99
 Db 241 GTGATCTCTTCGACCCCGCGCGCTGGCGCGGTCTGAACCAACCTGTTCGCAAGATG 300

Qy	100	LeuAsnGluMetLysGluTrpLysGluLysAspValSerValGlnLeuSerLeuLeuGly	119
Db	301	CTGGGGCAAGTGGCCCGTGGCAGCAAGGTCGCGAGTGCAGCATCGTGCACATCGGT	360
Qy	120	SerLysSerIleAsnPheGlnSerLeuGlyIleLysIleLeuThrGlnAspSerGly	139
Db	361	CAGAAAGCCTCGGGTCTCTCCGCCGAATCAAGTCAACATGTCGCGAGCGTCACGCAT	420
Qy	140	IleGlyAspThrProSerValGlnLeuLeuIleGlySerValAsnSerMetIleAspAla	159
Db	421	CTTGGCGACAGCCGCACATCGAGCAGTTGGTGGCGTGATCAAGGTGATGTGATGCC	480
Qy	160	TyrLysLysGlyGluValAspValValTyrLeuValTyrAsnLysPheIleAsnThrMet	179
Db	481	TTTACCGAGGGCAAGTGGACGGGTGTTATCTGTCTACAACCGCTTCGTGACACCATG	540
Qy	180	SerGlnLysProValLeuGluLysLeuIleProLeuProGluLeuAspAsnAspGluLeu	199
Db	541	ACGCAGAGGCCAGCTTCGACACATGTCGCGCTTCGG-----GCC	582
Qy	200	GlyGluArgLysGlnVal-----TyrAspTyrIleTyrGluProAspAlaLysVal	216
Db	583	GCCGAGCAAGGTGGCGCACCCAGCTGGGACTACTGTACGACCGCGTGCCTCCACC	642
Qy	217	LeuLeuAspAsnLeuLeuValArgTyrLeuGluSerGlnValTyrGlnAlaValGlu	236
Db	643	GTGCTGGAGCACGTGATGACGCGCTACATCAGTGTGCTGTACAGCGCGTGTGAA	702
Qy	237	AsnLeuAlaSerGluGlnAlaAlaArgMetValAlaMetLysAlaAlaThrAspAsnAla	256
Db	703	AACGTGGCCTCCGAGCAGCGCGCGCGCATGGTGCCATGAGCGCCGACGCGAACGCC	762
Qy	257	GlyAsnLeuIleAsnGluLeuGlnLeuValTyrAsnLysAlaArgGlnAlaSerIleThr	276
Db	763	AACAAGATGATCGCACCTTGCACCTGGTCTACACAAGCGCGCCGCGCGCATCC	822
Qy	277	AsnGluLeuAsnGluIleValAlaGlyAlaAlaAlaIle	289
Db	823	CAGAAATTTTCGAAATTCGTACCGCGCGCGCAGCCGTA	861

Search completed: November 9, 2004, 22:10:43
Job time : 2233 secs